UNIVERSITÉ PARIS-SUD

ÉCOLE DOCTORALE : Sciences et Technologie de l'Information, des Télécommunications et des Systèmes

Laboratoire des Signaux et Systèmes (L2S)

DISCIPLINE : Physique

THÈSE DE DOCTORAT

soutenance le 11 Décembre 2014

par

Benjamin LARROUSSE

Structure d'information, stratégies de communication et application aux réseaux distribués.

Composition du jury :

Directeur de thèse :	Samson LASAULCE	Directeur de Recherche CNRS (L2S)
Rapporteurs :	Olivier GOSSNER	Directeur de Recherche CNRS (Paris School of Economics)
	Rida LARAKI	Directeur de Recherche CNRS (Université Paris Dauphine)
Examinateurs :	Eitan ALTMAN	Directeur de Recherche INRIA (Sophia-Antipolis)
	Jean-Claude BELFIORE	Professeur (Telecom Paris-Tech)
	Matthieu BLOCH	Professeur associé (Georgia Institute of Technology)
	Michèle WIGGER	Professeur associé (Telecom Paris-Tech)

Remerciements

Je voudrais tout d'abord remercier mon encadrant *Samson Lasaulce* qui a su me convaincre de me lancer à fond dans cette aventure, qui m'a accueilli dans son équipe et sans qui cette thèse n'aurait pas été possible. J'ai eu le plaisir d'intégrer l'équipe alors composée de *François* et *Vineeth*, que je remercie pour leur accueil ainsi que pour tous les excellents moments passés ensemble au laboratoire, à échanger sur tout et n'importe quoi. La rapide intégration et la bonne ambiance du début de thèse n'auraient pas eu lieu non plus sans mes aînés doctorants *Elsa, Jean-François, José, Amadou, Lana, Zeina, Thang, Francesca*, ainsi que ceux qui sont arrivés à peu de choses près en même temps que moi : *Anna, Mathieu, Etienne, Pierre et Olivier*. Merci particulièrement aux trois derniers cités pour nos séances communes sportives indispensables pour être en forme pendant le marathon qu'est une thèse. Je n'oublie pas non plus les doctorants arrivés après moi et qui contribuent toujours à l'ambiance du laboratoire : *ChengFang, Achal, Diane, Florian, Chao*.

Je remercie également *Thierry Lafay* de m'avoir accepté dans son équipe d'enseignement à l'Université Paris I Panthéon-Sorbonne. L'ambiance d'équipe était très bonne, et enseigner fut très bénéfique pour ma thèse et aussi et surtout pour moi personnellement. Mes pensées vont également à *Matthieu Bloch* pour son apport conséquent lors de ma thèse.

Enfin, je remercie du fond du cœur mes parents, qui m'ont supporté pendant une bonne partie de la rédaction et pendant les vingt-cinq années qui ont précédé ; ma sœur *Hélène*, mon oncle *Nicolas*, et mes grands-parents qui m'apportent beaucoup malgré la distance qui nous sépare ; *Coline*, qui a été d'une infinie patience, qui a souffert avec moi ces deux dernières années et qui m'a apporté l'équilibre dont j'avais besoin ; ainsi que mes amis, pour leur soutien sans faille, leur joie de vivre et pour tout ce qu'ils m'apportent mais qu'il est beaucoup trop long de détailler ici : *Audrey, Fabien, Ferdinand, Hadrien, Justine, Kévin, Mylène, Nathalie*.
Vous avez tous, chacun à votre façon, contribué à la réussite de cette thèse.

« I like the scientific spirit - the holding off, the being sure but not too sure, the willingness to surrender ideas when the evidence is against them : this is ultimately fine - it always keeps the way beyond open - always gives life, thought, affection, the whole man, a chance to try over again after a mistake - after a wrong guess. »

– Walt Whitman, Walt Whitman's Camden Conversations.

Sommaire

Table des figures ix

Liste des tableaux xi

Résumé xiii

Abstract xiv

Notations xv

Chapitre 1
Introduction 1

1.1 Contexte de la thèse . . . 1
1.1.1 Généralités . . . 1
1.1.2 La théorie de l'information . . . 3
1.1.3 La théorie des jeux . . . 4
1.1.4 Le contrôle de puissance dans les réseaux sans fil . . . 7
1.1.5 Les réseaux de véhicules électriques . . . 8
1.2 Structure et objectifs du manuscrit . . . 8
1.2.1 Exemple introductif : Le « Big Online Matching Game » . . . 9
1.2.2 Exemple introductif : les cas de référence . . . 10
1.2.3 Structure de la thèse . . . 14
1.3 Contributions et publications . . . 14
1.3.1 Contributions . . . 14
1.3.2 Publications . . . 16

Chapitre 2
Coordination via la communication implicite dans les réseaux distribués. 17

2.1 Motivation et état de l'art . . . 17
2.1.1 Un problème d'optimisation distribuée . . . 17

2.1.2 État de l'art . . . 19
2.1.3 Contributions principales . . . 20
2.2 Description du problème étudié . . . 21
2.2.1 Formulation mathématique . . . 21
2.2.2 Discussion . . . 23
2.3 Contraintes d'information sur les distributions atteignables . . . 26
2.3.1 Une condition nécessaire d'atteignabilité . . . 26
2.3.2 Conditions suffisantes d'atteignabilité . . . 28
2.4 Problème d'optimisation sous contraintes . . . 32
2.4.1 Cas général . . . 32
2.4.2 Le cas de l'observation parfaite $Y = X_1$ . . . 35
2.5 Application : le *contrôle de puissance codé* . . . 37
2.5.1 Description du cadre proposé . . . 37
2.5.2 Influence de la fonction d'utilité . . . 40
2.5.3 Influence de la structure d'observation . . . 42
2.5.4 Influence de la connaissance sur l'état . . . 44
2.5.5 Influence du schéma de coordination . . . 44
2.6 Conclusion du chapitre . . . 48

Chapitre 3
Communication stratégique pour les smart grids **51**

3.1 Motivation et état de l'art . . . 52
3.1.1 Introduction . . . 52
3.1.2 État de l'art . . . 53
3.2 Modèle de jeu de signaux et exemple introductif . . . 54
3.2.1 Modèle de jeu de signaux . . . 54
3.2.2 Discussion . . . 56
3.2.3 Un exemple illustratif . . . 57
3.3 Résultats théoriques . . . 61
3.3.1 Le cas de la dimension 1 . . . 61
3.3.2 Discussion en dimension supérieure . . . 64
3.4 Application numérique . . . 64
3.4.1 Influence du biais b . . . 65
3.4.2 Exemple de partitions et actions associées . . . 66
3.5 Conclusion du chapitre . . . 67

Chapitre 4
Transformation d'une structure d'observation arbitraire à l'aide d'un encodeur 69

4.1 Motivation et état de l'art . . . 69
4.1.1 Introduction . . . 69
4.1.2 Etat de l'art . . . 71
4.2 Description du modèle . . . 72
4.2.1 Structure d'observation avec assistance d'un encodeur . . . 72
4.2.2 Source aux variations arbitraires et *observation parfaite virtuelle* . . . 74
4.2.3 Lien avec les jeux répétés . . . 75
4.3 Contraintes d'information pour retrouver l'observation parfaite . . . 76
4.3.1 Graphe auxiliaire et coloriage . . . 76
4.3.2 Condition suffisante pour l'*observation parfaite virtuelle* . . . 77
4.3.3 Condition nécessaire pour l'*observation parfaite virtuelle* . . . 79
4.4 Application aux jeux répétés . . . 80
4.5 Conclusion du chapitre . . . 81

Chapitre 5
Conclusion et perspectives 83

Annexes 85

Annexe A
Théorie de l'information 85

A.1 Définitions et propositions de base . . . 85
A.2 Théorèmes de codage . . . 88
A.3 Séquences typiques . . . 89
A.4 Lemmes pour les théorèmes de codage . . . 90
A.5 Deux modèles de canaux en théorie de l'information . . . 92

Annexe B
Théorie des jeux 93

B.1 Quelques définitions utiles . . . 93
B.2 Equilibre de Nash et théorème Folk . . . 95
B.3 Dynamique de meilleure réponse . . . 96

Annexe C
Démonstrations du chapitre 3 **97**

C.1 Preuve de la proposition 3.1 . . . 97
C.2 Preuve de la proposition 3.2 . . . 99
C.3 Preuve de la proposition 3.3 . . . 101
C.4 Preuve de la proposition 3.4 . . . 102
C.5 Preuve de la proposition 3.5 . . . 102
C.6 Preuve du théorème 3.7 . . . 103

Annexe D
Démonstrations du chapitre 4 **105**

D.1 Preuve du Théorème 4.7 . . . 105
D.2 Preuve du Théorème 4.8 . . . 111
D.3 Écriture de la contrainte d'information pour l'application considérée . . . 116

Annexe E
Publications **119**

E.1 Revues internationales . . . 119
E.1.1 IEEE-TIT 2014 . . . 120
E.2 Conférences internationales avec comité de lecture et actes . . . 151
E.2.1 IEEE-ISIT 2013 . . . 152
E.2.2 IEEE-WiOpt 2014 . . . 158
E.2.3 IEEE-ICASSP 2014 . . . 165

Bibliographie **171**

Table des figures

1.1 Pierre-feuille-ciseaux . . . 4
1.2 Cattedrale di Santa Maria del Fiore, Florence, Italie. . . . 7
1.3 Le « Big Online Matching Game » (BOMG) . . . 9
1.4 Le « BOMG » : communication pure . . . 10
1.5 Le « BOMG » : « Cheap talk » avec préférences non alignées . . . 11
1.6 Le « BOMG » : le cas « Online Matching Pennies » [46] . . . 12
1.7 Stratégie sans communication pour « Online Matching Pennies » . . . 12
1.8 Stratégie de communication « paire/impaire » . . . 13

2.1 Schéma de coordination simple . . . 26
2.2 Illustration du schéma de codage . . . 29
2.4 Canal avec retour d'information RSIB . . . 39
2.5 Gains relatifs, cas somme des efficacités énergétiques . . . 42
2.6 Gains relatifs, cas somme des taux et sommes des RSIB . . . 42
2.9 Influence de la structure d'observation . . . 44
2.10 Comparaison des cas I et II . . . 45
2.11 Un jeu de contrôle de puissance . . . 45
2.14 Utilité espérée vs RSB(dB) pour différents scénarios . . . 48

3.1 Schéma de signal pour un modèle de Smart Grids . . . 54
3.3 Partitions optimales de taille 5 . . . 61
3.4 Partitions optimales en fonction de la probabilité sur l'état . . . 61
3.5 Nombre de cellules utilisées . . . 65
3.6 Gains relatifs des agents . . . 66
3.7 Exemple de partition d'équilibre en dimension 2 . . . 67
3.8 Charges effectives versus biais (dimension 2) . . . 68

4.1 Structure d'observation avec encodeur . . . 73
4.3 Structure d'observation privée $\mathcal{T}$ . . . 80
4.4 Structure du canal de diffusion $\mathcal{T}_C$ . . . 81
4.5 Comparaison des régions d'utilités . . . 82

A.1 Canal de diffusion . . . 92
A.2 Canal à accès multiple . . . 92

Liste des tableaux

2.3 Indexation choisie 35
2.7 Distributions marginales et distributions conjointes optimales *sous* contraintes . . 43
2.8 Distributions marginales et distributions conjointes optimales *sans* contraintes . . 43
2.12 Code canal proposé 47
2.13 Codage et décodage source proposé 47

3.2 Taille maximale des partitions d'équilibre 60

4.2 Matrice de paiements du Dilemme du prisonnier. 80

B.1 Matrice de paiements pour le jeu Faucon et Colombe. 94

Résumé

Cette thèse étudie des problèmes d'optimisation distribuée avec différentes structures d'observations et leurs applications aux réseaux sans fil et aux problèmes de Smart Grids. Spécifiquement, une structure d'observation asymétrique entre deux agents est considérée, où un premier agent a connaissance complète à propos de la réalisation d'un état aléatoire, et l'autre agent ne connaît rien à propos de cet état. Dans ce contexte, la question est de savoir comment transmettre de l'information depuis le premier agent vers le second agent dans le but d'utiliser de manière optimale les ressources de communication. Plusieurs modèles sont étudiés dans cette thèse. Pour tous, un élément commun est le fait que la source d'information doit être encodée de manière appropriée pour optimiser l'utilisation de la configuration du système. Un premier modèle est étudié où aucun canal de communication n'est disponible entre les agents et ils ont une fonction d'utilité commune. Cependant, le seul moyen de communiquer est via les actions choisies par les agents. Comme les actions ont une influence sur le paiement, l'agent informé encode sa connaissance à propos de l'état dans ses actions, qui seront observées de manière imparfaite par le second agent. Ce dernier décodera l'information et choisira ses actions dans le but de maximiser la fonction objectif commune. Nous utilisons des outils de *théorie de l'information* pour caractériser ce compromis optimal par une *contrainte d'information*, et appliquons ce scénario à un problème de contrôle de puissance pour un canal à interférence. Notre nouvelle stratégie (le *contrôle de puissance codé*) donne des gains très prometteurs comparés aux approches classiques.

Dans une seconde partie, nous considérons qu'il existe un canal dédié de communication, c'est-à-dire que les actions de l'agent informé n'ont pas d'influence sur le paiement et sont seulement utiles pour la transmission d'information. De plus, les agents sont supposés avoir des intérêts divergents, si bien que l'agent informé n'a pas nécessairement d'incitation à envoyer tout son savoir à l'agent non informé. La théorie des jeux et les jeux de « Cheap talk » en particulier sont le bon cadre pour analyser ce genre de problème. Nous caractérisons le schéma de signal sur lequel les agents se seront mis d'accord. Ce schéma amènera à un *équilibre de Nash*, est donc optimisera la façon dont la communication est faite. Ce modèle est d'un intérêt particulier pour les réseaux de véhicules électriques où un véhicule électrique doit envoyer son besoin en terme de puissance de charge à un agrégateur qui choisira un niveau de charge effectif pour le véhicule électrique. Ce dernier ne se souciera que de son besoin, alors que l'agrégateur se soucie également de l'état du réseau. Ce modèle aide à optimiser la façon dont le réseau est utilisé.

Enfin, nous considérons un modèle avec plus de deux agents, où le but principal est pour tous les agents de retrouver l'observation parfaite des actions passées de tous les agents. Ceci est d'un intérêt très particulier d'un point de vue de la théorie des jeux pour caractériser les utilités espérées de long terme des agents. Dans ce modèle, nous ajoutons un encodeur qui observe parfaitement toutes les actions passées et aidera les agents à obtenir l'observation parfaite. En fait, ceci sera possible si la bonne *contrainte d'information* est satisfaite. Nous caractérisons donc cette dernière, en utilisant un schéma de codage hybride combinant des outils classiques de théorie de l'information ainsi que des outils de la théorie des graphes.

Mots-clés: Optimisation distribuée, Coordination, Théorème de Codage, Contraintes d'Information, Théorie des Jeux, Théorie de l'Information, Quantification, Contrôle de Puissance, Smart Grids.

Abstract

Information structure, communication strategies and application to distributed networks

This thesis studies distributed optimization problems with different observation structures and application to wireless network and Smart Grids problems. Specifically, an asymmetric observation structure between two agents is considered, where a first agent has full knowledge about the realization of a random state, and the other agent does not know anything about this state. In this context, the question is how to transmit information from the first agent to the second agent in order to use in an optimal way the communication resources. Several models are studied in this thesis. For all of them, a common element is that the information source has to be encoded in an appropriate manner to optimize the use of the system's configuration. A first model is studied where no dedicated channel for communication is available between agents and they have the same objective function. Therefore, the only way communication is possible is through the actions chosen by agents. As actions are payoff relevant, the first agent has to find the optimal tradeoff between transmission of information and payoff maximization. The informed agent encodes his knowledge about the state into his actions, which will be imperfectly observed by the second agent. The latter will decode the information and choose his actions in order to maximize the common objective function. We use tools from *information theory* to characterize this optimal tradeoff by an *information constraint*, and apply this scenario to a power control problem in an interference channel setting. Our new strategy (the *coded power control*) gives some promising gains compare to classical approaches.

In a second part, we consider that there exists a dedicated channel for communication, that is to say the actions of the informed agent are not payoff relevant and are only useful for transmission of information. Furthermore, agents are supposed to have diverging interests, so that the informed agent does not necessarily have an incentive to send all his knowledge to the uninformed agent. Game theory and Cheap talk game in particular appears to be the right framework to analyze this problem. We characterize the signal scheme that agents will agree on. This scheme will lead to a *Nash Equilibrium*, thus will optimize the way communication is done. This model is of particular interest for electrical vehicles networks where an electrical vehicle has to send his need in term of power to an aggregator which will choose an effective charging level for the electrical vehicle. The latter only cares about his need in term of power whereas the aggregator also takes into account the network status. The considered model help to optimize the way the network is used.

We finally consider a model with more than two agents, where the main goal is for all agents to retrieve perfect observations of all past actions of all agents. This is of particular interest in a game theory point of view to characterize the long term expected utilities of the agents. In this model, we add an encoder who perfectly oberves all past actions and will help agents to have perfect monitoring. In fact, this is possible if the right *information constraint* is satisfied. We thus characterized the latter, using a hybrid coding scheme combining classical information theoretic scheme and tools from graph theory.

Keywords: Distributed optimization, Coordination, Coding theorem, Information constraint, Game Theory, Information Theory, Signal quantization, Power Control, Smart Grids.

Notations

Symbole	Signification
$\mathbb{N}$	Le corps des nombres entiers
$\mathbb{R}$	Le corps des réels
$\mathbb{R}_+$	$[0, +\infty)$
$\mathbb{1}_{\{\cdot\}}$	Fonction indicatrice
$\oplus$	Addition modulo-2
$\otimes$	composée entre deux applications
$\triangleq$	Utilisé pour définir une nouvelle variable
Z	Variable aléatoire générique
Z_i^j	Séquence de variables aléatoires $(Z_i, ..., Z_j)$, pour deux entiers $i < j$
$\mathcal{Z}$	Alphabet de Z
$\lvert\mathcal{Z}\rvert$	Cardinalité de $\mathcal{Z}$
$\Delta(\mathcal{Z})$	Simplexe unité sur $\mathcal{Z}$
z	Réalisation de Z
z^n or $\underline{z}$	Séquence ou vecteur $(z_1, z_2, ..., z_n)$
$\underline{z}_{-i}$	Toutes les composantes du vecteur $\underline{z}$ sauf la $i$ème
$\mathbb{E}_{\mathrm{P}}$	Espérance mathématique par rapport à la distribution de probabilité P
$H(Z)$	Entropie de Z
$I(Y;Z)$	Information mutuelle entre Y et Z
$Z_1 - Z_2 - Z_3$	Chaine de Markov : $\mathrm{P}(z_1\lvert z_2, z_3) = \mathrm{P}(z_1\lvert z_2)$
$\mathcal{T}$	Probabilité de transition d'un canal
$\mathcal{T}_\epsilon^n(\cdot)$	Ensemble des séquences typiques (voir Définition A.17)

Chapitre 1

Introduction

« The difference between mathematicians and physicists is that after physicists prove a big result they think it is fantastic but after mathematicians prove a big result they think it is trivial. »

– Lucien Szpiro, mathématicien français, pendant un cours d'algèbre.

1.1 Contexte de la thèse

1.1.1 Généralités

De nos jours, quand nous évoquons la communication, il nous faut préciser à quoi nous faisons référence. D'une part, la communication peut être vue comme un moyen de transmettre des données, ce qui est un sujet très complexe et sophistiqué. Les moyens de communication modernes sont utilisés mondialement pour transmettre de manière quasi instantanée via la télévision, internet, ou les smartphones, des contenus très variés tels que les informations (souvent illustrées par des photos et des vidéos, ce qui peut faire beaucoup de données), des documents professionnels (de plus en plus de personnes travaillent à distance de nos jours), des films, de la musique, voire même des quantités insoupçonnées de photos de chats. Pour ce faire, et avec l'augmentation des flux de données, de nombreux outils théoriques et pratiques sont développés pour permettre ces échanges.
D'autre part, la communication peut être vue comme un « simple » échange entre deux personnes. C'est un moyen de transmettre des connaissances. Il y a ce besoin de compréhension mutuelle : si les personnes qui échangent ont un objectif commun, il est évident que communiquer peut être bénéfique pour l'atteindre. Dans le cas d'intérêts conflictuels, communiquer ses intentions peut également être utile d'un point de vue stratégique pour arriver à ses fins. Certains grands conflits historiques tels que la crise de Cuba (1962) n'auraient sans doute pas eu la même issue sans communication.
Que les objectifs des différents agents soient alignés ou non, l'absence de communication peut entraîner des conséquences désastreuses. Un des exemples les plus insolites, tiré de [37] est le suivant : sur une petite route de montagne en Italie, une petite section a été réduite à une voie. Les voitures arrivant de directions opposées passaient donc tour à tour via un accord implicite. Jusqu'à ce que deux voitures s'engagent en même temps et se bloquent mutuellement le passage. Refusant de reculer, et les voitures s'accumulant derrière, il fallut trois jours aux autorités italiennes pour faire disparaître cet embouteillage !

Communiquer est donc primordial dans bon nombre de situations, et peut prendre plusieurs formes :

- une forme *explicite*, comme par exemple fixer le lieu et la date d'un rendez-vous, ce qui de nos jours est grandement simplifié par les moyens de communication modernes, ou encore répartir les tâches dans un travail de groupe, et ce via un échange en face à face, téléphonique, de SMS ou par mail. Selon le contexte, ces deux exemples peuvent rendre compte d'objectifs communs ou non pour les agents : deux personnes qui ont pour objectif commun de passer du temps ensemble chercheront juste un accord sur l'heure et l'activité qu'ils feront [1] ; en revanche, dans une entreprise, chaque individu peut avoir un objectif personnel plus ou moins éloigné de l'objectif général de son entreprise. Dans ce cas, la communication peut amener à réduire partiellement ou complètement le biais entre les objectifs.
  Un autre exemple de communication explicite est l'action de freiner en voiture. En effet, l'appui sur la pédale de frein actionne instantanément les feux stop, qui signalent aux automobilistes situés derrière le véhicule que ce dernier freine et qu'il faut probablement freiner aussi pour éviter la collision. Bien que cet exemple fasse partie de la communication explicite, l'information « cette voiture est en train de freiner » n'est pas envoyée telle quelle, elle est *codée* en un message (les feux stop sont allumés). Comme nous le verrons dans ce manuscrit, coder l'information est primordial.
  Un dernier exemple de conflit où la communication a un rôle central est l'ultimatum. Nous pouvons citer l'un des plus célèbres, celui du 23 juillet 1914 [2] de l'Autriche-Hongrie à la Serbie, qui entraînera, par le jeu des diverses alliances militaires de l'époque, la Première Guerre mondiale. Cet ultimatum est un exemple où le résultat de l'échange aboutit à la meilleure solution pour aucune des parties ;

- une forme *implicite*. Nous l'utilisons quotidiennement sans même nous en rendre compte. Un exemple basique est le simple fait de marcher dans la rue. En effet, quand deux personnes se retrouvent face à face, faire un mouvement pour indiquer de quel côté la personne va éviter l'autre est une forme de communication *implicite* : l'échange n'est pas verbal, il se fait via un mouvement, une *action*. Dans ce cas précis, l'action est interprétée de manière claire par la personne en face, qui pourra à son tour choisir le « bon » côté, celui qui permet d'éviter la collision. Il est important de souligner ici que nous avons supposé que la communication explicite à voix haute « Je vous évite par ce côté ! » n'est pas une option, car dans les faits, personne ne l'envisage.
  La communication n'a pas toujours pour but d'améliorer la connaissance de celui qui reçoit le message. Parfois, c'est l'effet inverse et la communication amène à une plus grande incertitude sur l'information détenue par celui qui envoie le message, ce qui peut lui être bénéfique. Au football lorsqu'un tir au but s'apprête à être effectué, la position du gardien, la course d'élan du tireur ou encore la position de son pied font autant office d'actions que de messages, aboutissant à de la communication implicite. Par exemple, en 2002, Mickaël

1. Un modèle où deux agents ont des préférences différentes, c'est-à-dire où le choix du lieu et de l'activité importe, mais où ils préfèrent quand même être ensemble que seuls, est la « la Bataille des sexes », bien connue en théorie des jeux. Une version de ce modèle, auquel une phase de communication est ajoutée au préalable, est étudiée dans [9].
2. Pour plus de détails, le lecteur peut se référer à
http://fr.wikipedia.org/wiki/Ultimatum_du_23_juillet_1914 .

Landreau, alors gardien du FC Nantes, s'est positionné sur la gauche de son but [3] (et non au milieu comme habituellement), envoyant alors un message qui perturba le tireur brésilien du Paris Saint-Germain Ronaldinho. Ce qui fonctionna puisque Landreau arrêta le tir au but.
D'autres exemples peuvent se retrouver au poker, où le temps de réflexion et le montant des mises sont des actions ainsi que des messages aux adversaires.

Le choix de communiquer de manière explicite ou implicite n'est souvent pas libre et dépend très fortement du contexte et des possibilités de communication. Dans certaines situations, des moyens de communication explicite sont disponibles, rendant nul le coût de la communication : la signalisation du freinage a par exemple un coût nul. Dans d'autres situations, le seul moyen de communiquer est de manière implicite, ce qui peut avoir un coût, comme nous le verrons plus loin, notamment en Section 1.2 et au Chapitre 2.

L'un des points centraux de cette thèse est l'étude théorique de problèmes d'optimisation à préférence commune ou non [4] où la communication se fait via certaines structures d'information, ce qui nous amènera à caractériser la manière « optimale » de se coordonner. Ici, nous considérerons que l'information à transmettre est une connaissance privée sur l'état (aléatoire) du système en question. Comme nous le verrons, il ne suffit pas d'envoyer l'information telle quelle (le système considéré ne permet pas toujours de le faire), il faut parfois la compresser, la coder en un certain nombre de messages. Ce n'est que grâce à ce codage que la transmission d'information est utilisée de manière optimale. Pour ce faire, nous utiliserons des outils de la *théorie de l'information*, qui est l'objet de la section suivante.

1.1.2 La théorie de l'information

Nota Bene : les définitions de bases ainsi que les lemmes, propositions, et théorèmes de la théorie de l'information utiles pour cette thèse sont proposés en Annexe A.

La théorie de l'information modélise la communication au sens technique du terme. La signification du message que nous voulons transmettre n'a que peu d'importance. La sémantique est exclue de cette théorie, l'essentiel étant de transmettre des messages de manière fiable. Pour ce faire, la communication est modélisée par différentes composantes : sources d'information, émetteurs, récepteurs, canaux de transmission (possiblement bruités), relais, etc. Grâce à cette modélisation, nous sommes en mesure de connaître les capacités théoriques de transmission d'information d'un certain nombre de systèmes. Par exemple, nous savons que le taux minimal de compression d'une source est égale à l'entropie de la variable aléatoire correspondant à cette source (Proposition A.13), et qu'il existe un code qui atteint ce taux minimal. Nous connaissons aussi la capacité maximale d'un canal de transmission point à point (Proposition A.14). De nos jours, avec l'augmentation du nombre d'échanges d'information, il est très important de savoir comment transmettre une information de la manière la plus concise possible, tout en gardant une grande fiabilité (typiquement une probabilité d'erreur au décodage proche de zéro).
D'autre part, ce genre de résultats, ainsi que les techniques de codage de source et de codage canal, nous seront très utiles dans cette thèse, que ce soit pour la communication implicite ou

3. La vidéo ici : https://www.youtube.com/watch?v=2akX3cYLSEg.
4. Nous parlerons d'objectif unique et de multi-objectifs.

explicite. En effet, lorsque nous sommes dans une configuration où la communication est faite de manière explicite, l'information est codée, compressée en un certain nombre de messages qui seront envoyés au récepteur. Nous ferons ici des liens avec la *quantification*, lorsque le nombre de messages qui peuvent être transmis est bien plus petit que le nombre d'états de la source. Il nous faudra créer un codage de la source d'information pour utiliser le système de manière optimale. Dans le cas de la communication implicite, les actions prises par un agent deviennent des messages, qui seront interprétés par un autre agent. Les messages sont donc codés en actions. Le nombre d'actions pouvant être plus petit que le nombre de messages à transmettre, il y a un besoin de compression des messages de la source. Grâce à la théorie de l'information, nous pouvons savoir comment coder ces messages de manière optimale.

La théorie de l'information a en fait des liens avec beaucoup d'autres domaines [26] :

- La théorie des communications, notamment parce que la croyance[5] selon laquelle il était impossible d'envoyer de l'information à un taux positif avec une probabilité d'erreur négligeable, a été démentie par Shannon [79] ;
- La physique, tout particulièrement la seconde loi en thermodynamique ;
- Les mathématiques, notamment les statistiques et la théorie des probabilités ;
- L'économie, via la théorie des portefeuilles ainsi que le *critère de Kelly*, relatif aux systèmes de paris et d'investissements.

Ici, nous contribuons à la théorie de l'information en renforçant les liens de cette théorie avec la *théorie des jeux*, qui est l'objet de la section suivante.

1.1.3 La théorie des jeux

Nota Bene : les définitions de bases ainsi que les lemmes, propositions, et théorèmes de la théorie des jeux utiles pour cette thèse sont proposés en Annexe B.

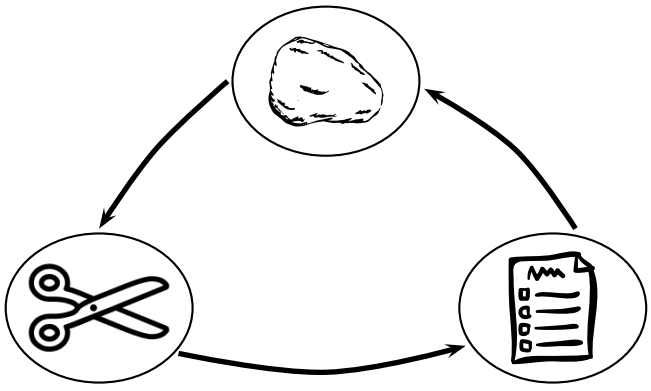

FIGURE 1.1 – **Pierre-feuille-ciseaux** : dans ce jeu, chaque agent a trois actions, selon le principe suivant : les ciseaux coupent le papier, le papier enveloppe la pierre, la pierre écrase les ciseaux. Grâce à la théorie des jeux, nous savons que la stratégie qui induit l'utilité espérée la plus élevée est de jouer chaque action avec probabilité $\frac{1}{3}$ (pour des agents rationnels).

5. Cette croyance est datée du début des années 1940.

La théorie des jeux est une branche des mathématiques introduite par John von Neumann et Oskar Morgenstern en 1944 [85], puis John Nash dans les années 1950 [72] et qui étudie des systèmes (le *jeu*) composés d'agents (les *joueurs*) qui interagissent. Chaque agent cherche à maximiser une utilité individuelle, qui dépend généralement des choix (appelés *actions*) de tous les agents du système. C'est une théorie de la décision pour des agents qui ont une influence les uns sur les autres et qui ont connaissance de ces influences réciproques. C'est pour cette raison qu'une telle situation est appelée *stratégique*, et les choix des agents sont des stratégies. Dans ce contexte, l'utilité d'un agent dépend de ses propres choix ainsi que des choix des autres agents. Cela amène des difficultés supplémentaires par rapport à de l'optimisation classique, où l'utilité ne dépend des choix que d'un seul agent. Bien qu'assez différente en termes d'approche, c'est une branche étroitement liée à l'optimisation. Par exemple, pour les jeux de potentiel[6] l'utilité maximale pour les agents correspond au maximum d'une fonction associée au jeu (la fonction potentielle). Un des concepts de solution (Nash [72]) est l'*équilibre de Nash* : les actions des agents forment un équilibre du jeu si aucun agent ne peut augmenter son utilité en choisissant une autre action pendant que les autres maintiennent leur choix d'actions d'équilibre. L'équilibre de Nash garantit une propriété de stabilité des choix des agents, sous l'hypothèse cruciale que les agents sont rationnels. Ce qui nous permet de caractériser à la fois le comportement des agents et également l'issue (les utilités effectivement obtenues par les agents) d'un jeu. Un exemple célèbre de « jeu » est le *Pierre, Feuille, Ciseaux* (Figure 1.1.3). Il est à noter qu'un équilibre de Nash, bien que stable, peut ne pas être efficace, c'est-à-dire qu'il n'engendre pas l'utilité maximale que les agents pourraient espérer.
Même si beaucoup de problèmes sont très complexes et certains non résolus à ce jour, au fil des années, cette théorie a évolué pour englober de nombreuses situations différentes (liste non exhaustive) :

- 1 agent, 2 agents, $n \in \mathbb{N}$ agents, et même une infinité d'agents ;
- Jeux à somme nulle (ce que l'un gagne, l'autre le perd) et somme non nulle ;
- Jeux répétés, pour lesquels nous répétons un jeu en un coup un nombre fini ou infini de fois, ce qui permet d'introduire de l'information supplémentaire (grâce à l'histoire du jeu) ;
- Information incomplète, lorsque chaque agent ne connait pas exactement ses propres possibilités d'action, les possibilités d'action des autres agents, les gains résultants des choix d'actions, ou bien les motivations des autres agents ;
- Observation imparfaite lorsque les actions passées des autres agents ne sont pas observées parfaitement (dans le cas d'un jeu répété) ;
- Jeux à mémoire parfaite, où chaque agent est capable à tout moment de se souvenir de la suite d'actions qui a été jouée précédemment, ou à mémoire imparfaite ;
- Jeux dynamique, stochastique.

Dès lors, cette théorie fournit un cadre adapté à la modélisation de nombreuses situations notamment économiques, pour preuve les différents *Prix Nobel en économie* qui sont théoriciens des jeux : John Nash, Reinhard Selten, et John Harsanyi (1994), Thomas Schelling et Robert Aumann (2005), Leonid Hurwicz, Eric Maskin et Roger Myerson (2007), et enfin Alvin Roth et Lloyd Shapley (2012).

6. Voir la définition ici : http://fr.wikipedia.org/wiki/Jeu_de_potentiel.

Depuis quelques années, avec l'essor des réseaux sans fil et le nombre gigantesque de données qui sont échangées chaque jour, de nombreux travaux ont vu le jour en vue d'appliquer des mécanismes de théorie des jeux pour des problèmes de communication. Une des premières formulations d'un problème de contrôle de puissance grâce à la théorie des jeux peut être trouvée dans [44]. Nous pouvons citer également un des livres de référence sur le sujet [58]. Un des avantages de ce domaine provient du fait que les différents éléments d'un système (téléphones portables, station de base, box, etc.) seront programmés grâce aux solutions théoriques trouvées, ce qui permet de satisfaire à l'hypothèse cruciale de rationalité des agents. Supposer dans un contexte économique que les agents agissent de manière rationnelle est plus difficile à justifier.

Une des extensions du concept d'équilibre de Nash est l'équilibre pour les jeux de communication. [56] est un article de référence sur le sujet. Le modèle de base est celui de Crawford et Sobel [27], où un agent a de l'information sur un état de nature qu'un autre agent n'a pas, il envoie donc un message vers cet agent non informé, ce dernier faisant un choix d'action. Les utilités des agents dépendent de l'état de la nature et de l'action du second agent, mais pas *directement* du message envoyé par le premier agent. La communication est gratuite, le choix d'un message n'a pas d'influence *directe* sur l'utilité des agents, c'est pourquoi nous parlerons de jeux de « Cheap talk »[7]. Nous nous intéresserons particulièrement à ce modèle dans le Chapitre 3. En fonction des hypothèses que nous faisons sur la communication (unilatérale/bilatérale, sans coût/avec coût, mono-phase/multi-phase, observation parfaite/imparfaite message certifiable/non-certifiable etc.), l'issue du jeu peut être radicalement différente (voir par exemple [24, 4, 65, 38, 19, 8]).
La structure de communication (donc de transmission d'information) entre les agents a donc une grande influence sur le résultat d'un jeu.

Les théorèmes parmi les plus importants en théorie des jeux pour les modèles de jeux répétés sont les théorèmes Folk (voir le théorème B.34 en annexe). Un théorème Folk caractérise les utilités d'équilibre d'un jeu répété de long terme (c'est-à-dire lorsque le nombre d'étapes tend vers l'infini). Une des hypothèses cruciales pour avoir un théorème Folk est celle d'*observation parfaite* des actions des autres agents. Cette hypothèse permet d'assurer aux agents la possibilité de se coordonner pour adopter des stratégies amenant à un équilibre de Nash. En effet, pour caractériser les équilibres de Nash, des stratégies « grim trigger » sont utilisées : chaque agent coopère, mais dès lors qu'un agent dévie de la stratégie coopérative, une action de punition est mise en place jusqu'à la fin du jeu. Pour pouvoir utiliser ce genre de stratégies, il est important de pouvoir détecter une déviation d'un agent. C'est ce que permet l'hypothèse d'observation parfaite des actions passées.
Cette hypothèse est l'objet central du Chapitre 4. Dans ce dernier, nous étudions, dans un système donné où nous supposons une structure d'observation arbitraire, comment concevoir un encodeur (qui sera ajouté au système) qui permettra de retrouver l'observation parfaite pour tous les agents.

La théorie des jeux, ainsi que la théorie de l'information, modélisent (entre autres) la communication entre plusieurs agents, avec deux approches très différentes mais que nous pensons complémentaires. La théorie des jeux permet de modéliser un système multi-agents avec préférences alignées ou non, d'aider à la décision pour avoir un résultat parfois optimal et en tout cas stable. Quant à la théorie de l'information, elle apporte des outils techniques pour calculer

7. Ces jeux sont également appelés jeux de signaux, ou encore « Sender-receiver games ».

les capacités limites de ces mêmes systèmes, en fonction de la structure d'information imposée. Dans cette thèse, nous utilisons les deux théories en montrant que des approches de l'une peuvent servir à l'autre et vice versa. Renforcer les liens entre théorie de l'information et théorie des jeux est un des objectifs de cette thèse. De plus, ces modèles seront appliqués à des problématiques de réseaux sans fil distribués ainsi qu'aux réseaux de véhicules électriques, que nous introduisons dans les deux sections suivantes.

1.1.4 Le contrôle de puissance dans les réseaux sans fil

(a) Le Duomo de l'extérieur

(b) Le Plafond du Duomo

FIGURE 1.2 – Cattedrale di Santa Maria del Fiore, Florence, Italie.

Florence, Cattedrale di Santa Maria del Fiore (Photo 1.2(a)). Les touristes s'émerveillent devant le plafond du Dôme (Photo 1.2(b)). Chacun chuchote à son voisin à quel point il trouve cela magnifique. Problème : si tout le monde commence à parler, cela crée des interférences, amenant tout le monde à parler plus fort pour se faire entendre, ce qui crée encore plus d'interférences. Si rien n'est fait, tout le monde parle très fort (au maximum de ses capacités), et le silence du lieu n'est pas respecté.
Pour remédier au problème, un message est diffusé dans la Cathédrale pour signifier aux touristes de réduire considérablement le volume et respecter ainsi le lieu. La recommandation est tout à fait suivie par les visiteurs.

Cette situation introductive est un exemple de contrôle de puissance. Dans un réseau de communication mobile, la puissance d'émission des téléphones portables doit être contrôlée afin d'émettre suffisamment fort pour que les données (les appels, sms, ou les données d'une page web peuvent être vu comme des flux de données) soient transmises sans erreur, et suffisamment faible pour ne pas créer trop d'interférences aux autres utilisateurs du réseau. Dans notre exemple, le contrôle est fait de manière centralisée : une entité (les gérants du Duomo) dicte le comportement à suivre à tous les agents (les touristes), qui sont priés de respecter cette recommandation.
Le contrôle n'est cependant pas toujours fait de manière centralisée. Dans les réseaux de communication sans fil, une gestion centralisée peut demander beaucoup de ressources, ne serait-ce que pour récolter les informations du réseau entier. De plus, effectuer des calculs de manière centralisée peut poser de gros problèmes de complexité. Il y a donc un besoin de décentraliser la gestion et de gérer l'information de manière locale. Ce type d'approche décentralisée pour le contrôle de puissance est par exemple étudié dans [67].

Les agents ont donc une certaine autonomie de décision, rendant le contrôle décentralisé. Chaque agent a une certaine information (privée) à propos du réseau et va s'en servir pour prendre une décision (par exemple une puissance d'émission). La transmission d'information dans ces réseaux pour satisfaire aux objectifs individuels des agents ou un objectif global pour le réseau rentre parfaitement dans le cadre d'étude de la théorie des jeux ainsi que de la théorie de l'information. C'est donc un sujet d'étude très important et c'est une application que nous allons considérer dans cette thèse.

1.1.5 Les réseaux de véhicules électriques

La transition énergétique que nous connaissons actuellement amène de nouvelles problématiques dans beaucoup de domaines et notamment les réseaux de transports et de distributions d'électricité à moyenne et basse tensions. Il est indispensable de les gérer de manière efficace et de nouvelles méthodes sont ainsi à l'étude pour adapter de manière plus intelligente l'offre et la demande sur ces réseaux : ces derniers sont alors appelés Smart Grids [8]. En effet, la consommation énergétique doit être optimisée de façon à satisfaire la demande mais aussi les contraintes réseaux telles que le vieillissement des transformateurs ou la minimisation des pertes Joule générées. Une gestion de l'énergie via les Smart grids permettra d'atteindre une meilleure efficacité énergétique, notamment en évitant la surproduction ainsi que les consommations superflues.
Particulièrement, les problèmes de communication entre opérateurs du système électrique et consommateurs nous ont intéressés dans cette thèse. Supposons qu'un « opérateur réseau » et un « opérateur de consommation » (un gestionnaire de parc de véhicules électriques ayant une certaine demande en termes d'énergie pour charger les véhicules sous sa responsabilité, ou encore un opérateur de bâtiments de type tertiaire) discutent ensemble de leurs futurs échanges et de la façon dont ces échanges d'information (en terme de besoin énergétique, de capacité de productions, etc.). Ces discussions préalables sont indispensables pour optimiser la façon d'utiliser le réseau. Au Chapitre 3, nous nous intéressons particulièrement à ces échanges en supposant qu'un véhicule électrique pourra envoyer sa demande en terme d'énergie à un opérateur, qui lui fournira un niveau de charge. Ce sont ces phases de choix que nous allons étudier en définissant des stratégies optimales en amont, et qui seront implémentées en temps réel pour réaliser effectivement ces phases de communication. Grâce à des modèles de théorie des jeux, que nous appliquons pour la première fois à des problématiques de Smart Grids, il nous sera possible de concevoir une stratégie de communication efficace.

1.2 Structure et objectifs du manuscrit

Dans cette section, nous introduisons un jeu simple de communication que nous avons créé et qui servira de fil conducteur tout au long de ce manuscrit. Ce jeu permet d'illustrer la communication implicite ainsi que la communication explicite. En effet, cet exemple regroupe, de manière simplifiée, les différentes *structures d'information* que nous avons étudiées dans cette thèse. Cet exemple permet donc d'expliquer les intuitions derrière les résultats des chapitres 2, 3 et 4.

8. Une rapide introduction au Smart Grid se trouve à l'adresse suivante : http://fr.wikipedia.org/wiki/Smart_grid.

1.2.1 Exemple introductif : Le « Big Online Matching Game »

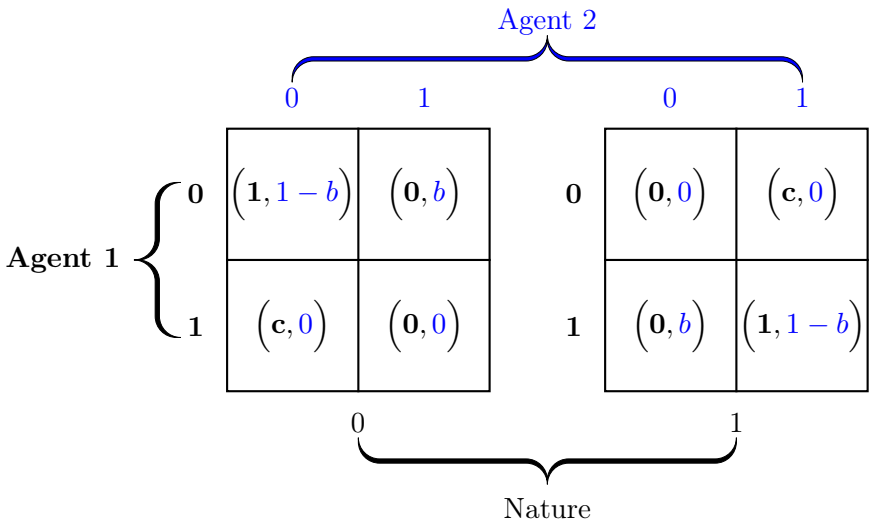

FIGURE 1.3 – Le « Big Online Matching Game » (BOMG) : la nature choisit la matrice de gauche avec probabilité $\rho_0 \in [0,1]$ et la matrice de droite avec probabilité $1-\rho_0$. L'Agent 1 choisit une ligne, l'Agent 2 une colonne. Ce jeu a deux paramètres : $b \in \mathbb{R}$ représente le « biais » entre les agents, et $c \in \mathbb{R}$ représente le coût de la communication.

Considérons le jeu sous forme matricielle[9] à deux agents avec état aléatoire de la nature représenté en Figure 1.3. Les agents ont deux actions possibles, dénotées "0" et "1", et la nature peut être dans deux états, également appelés "0" et "1". Considérons la répétition de ce jeu T fois, $T \in \mathbb{N}$. Les agents veulent maximiser leur utilité moyenne sur T étapes. A chaque étape, la nature choisit la matrice de gauche avec probabilité $\rho_0 \in [0,1]$ et la matrice de droite avec probabilité $1-\rho_0$. Nous supposons ici qu'un des agents (l'Agent 1) connaît toutes les réalisations de la nature en avance, et que l'autre agent (l'Agent 2) ne connaît rien sur ces réalisations. Ceci est la *structure d'information* de ce jeu. Cette structure imposera une contrainte sur les possibilités de communication entre les agents, que nous appellerons *contrainte d'information*. Dans ce jeu, il n'y a pas de canal explicite de communication, si l'Agent 1 a un intérêt à communiquer de l'information à l'Agent 2, il doit utiliser ses actions, c'est-à-dire faire de la *communication implicite*. Nous supposerons ici que l'Agent 2 **observe parfaitement** les états passés de la nature ainsi que les actions passées de l'Agent 1. De plus, ce jeu a deux paramètres :

- $b \in \mathbb{R}$ représente le « biais » entre les agents[10]

  → Si $b = \frac{1}{2}$, l'Agent 2 est indifférent entre jouer l'action 0 ou 1 : il obtiendra la même utilité dans les deux cas, quelque soit l'état de nature, et quelque soit le choix de l'Agent 1 ;

  → Si $b < \frac{1}{2}$, et l'état de nature est à 0, l'Agent 2 a une utilité plus élevée s'il joue l'action 0 que l'action 1 (quelque soit le choix de l'Agent 1). L'Agent 1 préfère lui aussi que l'Agent 2 joue l'action 0 : leurs préférences sont donc alignées. Pour l'état de nature 1, le même genre de raisonnement s'applique et les préférences sont aussi alignées ;

9. Pour la définition d'un jeu sous forme matricielle, voir l'exemple B.29 en annexe.
10. Cette analyse est vraie quelque soit la valeur du paramètre c.

→ Si $b > \frac{1}{2}$, et l'état de nature est à 0, l'Agent 2 a cette fois une utilité plus élevée en choisissant l'action 1 (quelque soit le choix de l'Agent 1). Or, s'il fait ce choix, l'Agent 1 aura lui une utilité plus faible (qu'il choisisse 0 ou 1) que si l'Agent 2 avait choisi l'action 0. Leurs préférences ne sont donc pas alignées.

- $c \in \mathbb{R}$ qui représente le coût de la communication :

  → Si $c = 1$, l'Agent 1 est indifférent entre l'action 0 et l'action 1, quelque soit l'état de la nature et le choix de l'Agent 2. Toutefois, il aimerait que l'Agent 2 joue l'action 0 lorsque l'état de nature est 0 et l'action 1 lorsque l'état de nature est 1. Il va donc utiliser ses actions pour transmettre les futures réalisations de l'état de la nature à l'Agent 2. Cela n'aura pas de coût pour lui, grâce à son indifférence entre ses deux actions ;

  → Si $c \neq 1$, l'Agent 1 a cette fois une préférence entre l'action 0 et l'action 1, qui dépend de la réalisation de la nature ainsi que du choix de l'Agent 2 : si $c < 1$, l'Agent 1 préfère l'action 0 pour l'état de nature 0 et l'action 1 pour l'état de nature 1, si $c > 1$ il préfère l'action 1 pour l'état de nature 0 et l'action 0 pour l'état de nature 1. Dès lors, utiliser ses actions pour communiquer engendrera un coût sur l'utilité moyenne finale. Malgré ce coût, la communication (implicite) peut être bénéfique et amener une utilité moyenne plus élevée par rapport à une stratégie sans communication.

1.2.2 Exemple introductif : les cas de référence

Pour ce jeu, il existe quatre cas à distinguer :

Cas 1) Biais $b \leq \frac{1}{2}$, coût $c = 1$: c'est le cas le plus simple. Il est représenté en Figure 1.4.

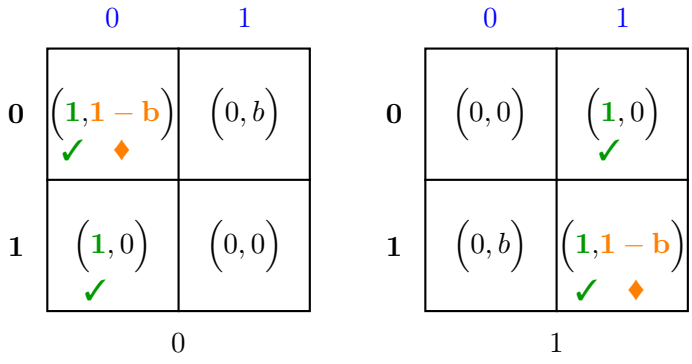

FIGURE 1.4 – **Cas 1)** : le « Big Online Matching Game », communication pure. Valeurs des paramètres : $b \leq \frac{1}{2}$ et $c = 1$. Les issues préférées pour chaque état de la nature de l'Agent 1 sont en vert et marquées du signe ✓, celles de l'Agent 2 en orange avec le signe ♦.

L'Agent 1 est indifférent entre l'action 0 et l'action 1, et peut donc utiliser ses actions pour communiquer. Le coût de la communication est nul et nous sommes dans un jeu de « Cheap Talk » avec préférences alignées, comme expliqué précédemment. L'agent 1 fait face à un problème de communication *pure* : ses choix d'actions n'ont pas d'influence (directe) sur son utilité. Il aimerait que l'Agent 2 joue 0 dans l'état 0 et 1 dans l'état 1, il a donc intérêt à échanger le maximum d'information possible. L'Agent 2 suivra la recommandation

de l'Agent 1 car il sait que c'est dans son intérêt. Une possibilité pour l'Agent 1 est de jouer à chaque étape l'état de la nature de l'étape suivante. Ainsi, l'Agent 2, qui observe parfaitement les actions passées de l'Agent 1, pourra jouer l'action optimale sur toutes les étapes sauf la première, et l'utilité sera maximale (sauf sur la première étape donc), ce qui donnera une utilité moyenne proche du maximum possible pour chacun des agents. Il est important de comprendre que les stratégies des agents sont décidées au préalable, donc que l'Agent 2 sait effectivement comment décoder l'information envoyée par l'Agent 1.

Cas 2) Biais $b \geq \frac{1}{2}$, coût $c = 1$: ce cas est représenté en Figure 1.5.

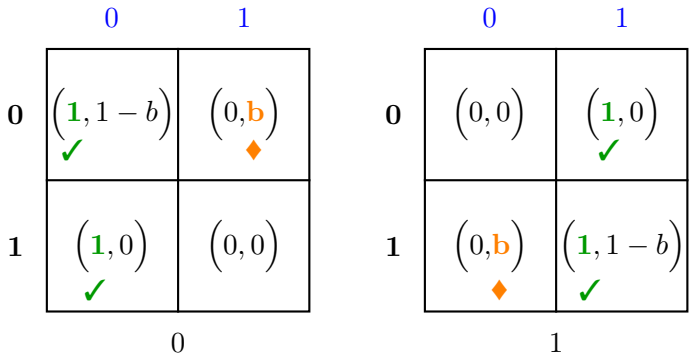

FIGURE 1.5 – **Cas 2)** : le « Big Online Matching Game », avec préférences non alignées. Valeurs des paramètres : $b \geq \frac{1}{2}$ et $c = 1$. Les issues préférées pour chaque état de la nature de l'Agent 1 sont en vert et marquées du signe ✓, celles de l'Agent 2 en orange avec le signe ♦.

Ce cas est similaire au cas 1), à l'exception près que les préférences des deux agents ne sont pas alignées. Dans ce cas, la communication n'a en fait pas d'intérêt : si l'Agent 1 signale l'état futur à l'Agent 2, il ne choisira pas l'action optimale du point de vue de l'Agent 1. Le mieux que l'Agent 1 puisse faire est donc de ne rien signaler. Cet exemple est en fait trop simple car il n'y a que deux actions disponibles et deux états. En revanche, si les nombres d'actions et d'états étaient très grands ou même infinis (si par exemple l'espace des possibilités était l'intervalle $[0, 1]$), atteindre un point d'équilibre nécessiterait d'envoyer une partie de l'information comme nous le verrons au Chapitre 3.

Cas 3) Biais $b \leq \frac{1}{2}$, coût $c \neq 1$: ici, les préférences sont alignées, mais le coût de la communication n'est pas nul. Si nous nous restreignons au cas $b = 0$ et $c = 0$ (Figure 1.6), nous sommes dans l'exemple « Online Matching Pennies » tiré de [46].

Ce cas est étudié en détail dans [46] et s'appelle « Online Matching Pennies ». Le but est que les deux agents se coordonnent avec l'état de nature, pour que les trois soient simultanément à 0 ou à 1. L'utilité moyenne maximale pour les 2 agents est ici de **1**. Un des résultats est que les stratégies de communication implicite peuvent induire une utilité moyenne bien plus élevée que des stratégies sans communication. Supposons que l'état de la nature soit 0 avec probabilité $\frac{1}{2}$ et 1 avec probabilité $\frac{1}{2}$, ce qui correspond au cas où

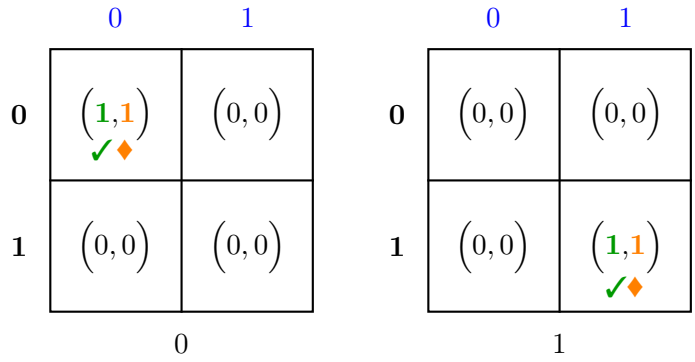

FIGURE 1.6 – **Cas 3)** : le « Big Online Matching Game », l'exemple « Online Matching Pennies » [46]. Valeurs des paramètres : $b = 0$ et $c = 0$. Les issues préférées pour chaque état de la nature de l'Agent 1 sont en vert et marquées du signe ✓, celles de l'Agent 2 en orange avec le signe ♦.

l'incertitude sur l'état de nature est maximale[11]. De simples stratégies montrent déjà les bénéfices de la communication via les actions :

→ Considérons tout d'abord une stratégie *sans communication*. Supposons pour cette première stratégie que l'Agent 1 joue exactement l'état de nature à chaque étape. Il ne transmet donc pas d'information à l'Agent 2 sur les futurs états de nature. L'Agent 2 jouera donc de manière aléatoire, par exemple il jouera l'action 0 et l'action 1 avec probabilité $\frac{1}{2}$. Cette stratégie est représentée en Figure 1.7.

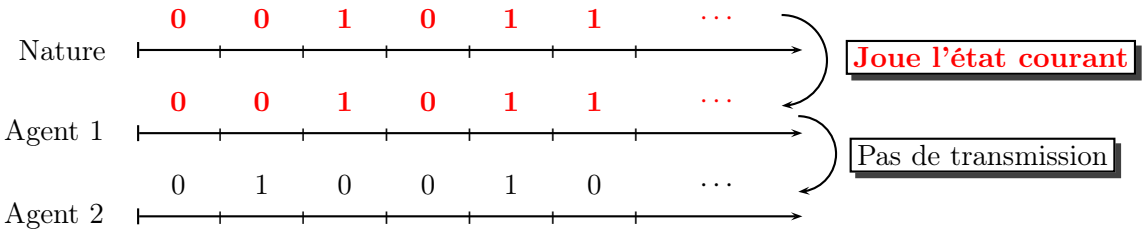

FIGURE 1.7 – Stratégie sans communication pour « Online Matching Pennies » : l'Agent 1 joue l'état courant à chaque étape, et ne transmet donc pas d'information. L'Agent 2 joue aléatoirement.

Avec cette stratégie, l'utilité moyenne espérée est de **0.5**.

→ Une seconde stratégie, avec transmission d'information cette fois, est pour l'Agent 1 de transmettre l'état de l'étape suivante aux étapes impaires et de jouer cet état aux étapes paires. Aux étapes paires, l'Agent 2 est donc informé de la réalisation de l'état de la nature et peut jouer cet état. Aux étapes impaires, il n'a pas d'information et joue aléatoirement. Cette stratégie est représentée en Figure 1.8.

11. Nous entendons par là que l'entropie de la variable aléatoire correspondant à l'état de nature est maximale.

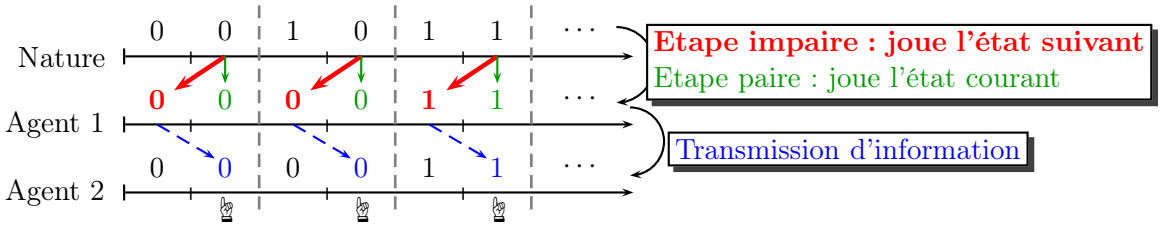

FIGURE 1.8 – Stratégie de communication « paire/impaire » : l'Agent 1 transmet de l'information sur la moitié des étapes (les étapes impaires) et joue de manière optimale sur l'autre moitié des étapes (les étapes paires). De cette façon, l'Agent 2 peut jouer de manière optimale sur la moitié des étapes (les étapes paires), et joue aléatoirement sur les autres (les étapes impaires). Cette stratégie sera utilisée pour le contrôle de puissance : voir Figure 2.1.

Avec cette stratégie, l'utilité moyenne est de **0.625**. Nous voyons donc que même avec une stratégie de communication très simple, l'utilité moyenne est bien plus élevée que celle correspondant à la stratégie sans communication.

→ L'analyse de [46] montre qu'il existe, toujours dans le cas d'*observation parfaite* des choix de l'Agent 1 par l'Agent 2, une stratégie de communication (lorsque le nombre d'étapes tend vers l'infini) qui induit une utilité moyenne de **0.81**. Cette valeur est la plus grande pouvant être atteinte, compte tenu de la structure d'information imposée.

Cet exemple sert de base à la communication implicite. Dans cet article, les auteurs caractérisent, pour tous jeux ayant la même *structure d'information* (en particulier l'observation parfaite et les préférences alignées), les utilités moyennes possibles que les agents peuvent atteindre. Dans cette thèse, nous nous servons de cet article comme point de départ, et nous étendons ces résultats à deux structures d'information plus générales, qui sont présentées au Chapitre 2 et en particulier en Section 2.2.

Cas 4) **Biais** $b > \frac{1}{2}$ **coût** $c \neq 1$: ici les agents ont des préférences non alignées et le coût de la communication n'est pas nul (voir Figure 1.3). Comme dans le cas 2), il n'y aura ici pas de communication. Pour des systèmes plus élaborés néanmoins, la communication peut être possible. Ce cas rentre dans l'analyse faite au Chapitre 2 bien que les préférences des agents ne soient pas alignées. L'ensemble des utilités moyennes pouvant être atteintes par les agents est caractérisée de la même manière, cependant dans ce cas nous ne chercherons pas à atteindre l'utilité moyenne maximale car chaque agent a ici sa propre utilité. La notion de maximum pour deux utilités distinctes n'a donc pas de sens, et il faut se tourner vers la notion d'équilibre de Nash.

Une des hypothèses cruciales dans cet exemple est celle d'*observation parfaite* des actions passées des autres agents. Dans le Chapitre 4, en plus de considérer un système avec un nombre d'agents supérieur à 2, nous étudions une structure d'observation *arbitraire* et explicitons la condition avec laquelle nous pouvons retrouver cette observation parfaite pour tous les agents. Cette condition sera également une *contrainte d'information*.

Maintenant que le contexte de cette thèse est expliqué, la section suivante présente la structure de ce manuscrit.

1.2.3 Structure de la thèse

La structure du manuscrit est articulée autour de l'exemple introduit en Sections 1.2.1 et 1.2.2. Elle est composée d'un chapitre d'introduction, de trois chapitres de contributions, chacun correspondant à un problème précis construit et motivé autour du « Big Online Matching Game », ainsi que d'un chapitre de conclusion et perspectives. Dans un souci de clarté, les principales définitions et propositions utiles pour cette thèse provenant de la théorie de l'information et de la théorie des jeux sont proposés en annexes A et B respectivement :

- L'Annexe A fournit les définitions, lemmes, et propositions fondamentaux de la théorie de l'information. En particulier, les lemmes utiles pour les schémas de codage de source et de codage canal sont présentés ;
- L'Annexe B fournit les définitions, lemmes, et propositions fondamentaux de la théorie des jeux. En particulier, nous mettons l'accent sur les résultats concernant l'équilibre de Nash, les jeux répétés ainsi que les théorèmes Folk, qui nous ont été utiles tout au long de cette thèse.

Les articles correspondants aux différentes contributions sont également proposés en Annexe E, ce qui permet notamment aux lecteurs de consulter les preuves qui ne seraient pas détaillées dans les chapitres 2 et 3.
Les démonstrations concernant le Chapitre 4 se trouvent en Annexe D.

1.3 Contributions et publications

Tout au long de ces trois années de thèse, nous nous sommes intéressés à des problèmes d'optimisation où des agents cherchent à atteindre un point d'entente, à se coordonner. Le terme « coordonner » est à prendre avec précaution ici : en effet, lorsque nous considérons des systèmes où tous les agents ont la même utilité, il est vrai de dire qu'ils cherchent à combiner leurs actions pour atteindre le meilleur résultat possible. Cependant, lorsque les agents ont des utilités divergentes, ils cherchent à obtenir l'utilité la plus élevée possible en tenant compte des choix d'actions des autres agents (un équilibre de Nash), les actions ayant une influence sur l'utilité. Ces choix ne vont pas forcément dans le même sens. Néanmoins, dans un souci d'uniformité, nous dirons dans ce cas que les agents cherchent à se coordonner sur un point d'équilibre.
Le thème commun des différentes parties de cette thèse est l'influence d'une certaine **structure d'information** associée à un système. En particulier, nous allons considérer des structures d'information **asymétriques**, où un agent sera mieux informé qu'un autre agent. Lorsque nous allons optimiser (au sens propre du terme) un système, la structure d'information imposera une **contrainte d'information**, qui variera selon le contexte considéré, et dont nous donnerons une caractérisation grâce aux outils de la théorie de l'information. Notamment, une caractéristique commune aux trois parties de cette thèse est la présence de **codage de source** pour envoyer de l'information d'un agent (informé) vers un autre (non informé ou partiellement informé).

1.3.1 Contributions

Les différentes contributions sont les suivantes :

- au Chapitre 2, nous considérons un problème d'optimisation distribuée où il n'y a pas de canal dédié à la communication, comme par exemple dans le cas 3) (Section 1.2.2) du

« Big Online Matching Game » . Ici, il y a deux agents, un agent informé par avance de toutes les réalisations d'un état de nature, et un autre agent qui ne connaît rien. Les agents maximisent la même utilité moyenne, le premier agent va donc utiliser ses actions pour transmettre de l'information au deuxième agent, tout en essayant de choisir des actions qui maximisent l'utilité moyenne. L'agent aura donc un compromis à trouver entre la transmission d'information et la maximisation de l'utilité moyenne. Ce modèle a été introduit dans [46] et [32] dans le cas où l'agent non informé observe parfaitement les actions passées de l'agent informé. Nous généralisons ces résultats en relâchant plusieurs hypothèses : nous considérons tout d'abord que l'agent non informé n'observe qu'une version bruitée des actions passées de l'agent informé. De plus, nous considérons le cas où l'agent non informé a une connaissance strictement causale de l'état ainsi que le cas où il n'a aucune observation de cet état. En plus de la caractérisation de la *contrainte d'information* (Section 2.3), nous étudions en profondeur le problème d'optimisation associé (Section 2.4), et nous faisons une application détaillée à un problème de contrôle de puissance dans les réseaux sans fil (Section 2.5) ;

- au Chapitre 3, nous relâchons l'hypothèse sur le coût de communication, et nous supposons maintenant que les actions de l'agent informé n'influent pas (directement) sur l'utilité des agents. En revanche, nous considérons un cadre plus général pour les utilités : ici, les agents ont des utilités distinctes, comme dans le cas 2) (Section 1.2.2) de l'exemple introductif. Nous passons donc de l'optimisation distribuée à la théorie des jeux, et en particulier aux jeux de « Cheap Talk », que nous allons appliquer pour la première fois à une problématique de Smart Grids. Le but est de concevoir des systèmes de signalisation entre un consommateur et un fournisseur d'électricité en amont, et qui seront implémentés en temps réel pour effectivement réaliser ces échanges d'information. Le consommateur veut atteindre son besoin en puissance, et envoie une requête (un message) au fournisseur, qui ne correspond pas forcément à son réel besoin. Le fournisseur qui reçoit cette requête veut à la fois satisfaire la demande tout en prenant en compte les coûts imputés au réseau. Nous explicitons et analysons ici les stratégies d'équilibre des deux agents.
  Nous faisons également des connexions entre le cadre étudié et le procédé de quantification en traitement du signal.

- au Chapitre 4, nous augmentons cette fois le nombre d'agents et nous considérons un nombre quelconque d'agents. Nous étudions également des jeux avec une *structure d'information arbitraire*. L'hypothèse d'observation parfaite n'est pas vérifiée, les agents reçoivent simplement un signal privé (sur les actions passées de tous les agents) via une certaine structure d'observation. Nous supposons également que la source d'information (les actions passées) peut varier d'étape en étape, nous considérons donc une source aux variations arbitraires (voir la Définition 4.1). Notre but est d'ajouter un encodeur qui aidera à rétablir l'hypothèse d'observation parfaite. Le cas où un signal public est envoyé à tous les agents est traité dans [60]. Ici, nous étendons ce résultat en ne supposant plus que le signal est public, mais que chaque agent reçoit un signal privé : nous ajoutons un canal de diffusion [12] avec information adjacente (voir Section A.5). Nous exhibons une *contrainte d'information* qui permet de garantir l'observation parfaite pour tous les agents, hypothèse très importante en théorie des jeux pour garantir l'existence d'un théorème Folk. Pour expliciter cette

12. « Broadcast channel » en anglais.

contrainte d'information, nous aurons besoin d'utiliser des outils de la théorie des graphes.

1.3.2 Publications

Cette thèse a donné lieu à des publications dans des conférences internationales ainsi que des revues internationales. Ces articles peuvent être trouvés dans l'Annexe E.

Articles dans des revues internationales

Le premier article est lié aux contributions présentées au Chapitre 2, tandis que le deuxième est lié à celles présentées au Chapitre 3 :

- B. Larrousse, S. Lasaulce, and M. Bloch, *Coordination in distributed networks. Application to power control*, soumis à IEEE Transaction on Information Theory ;
- O. Beaude, B. Larrousse, and S. Lasaulce, *Strategic communications in smart grids*, bientôt soumis à IEEE Transactions on Signal Processing.

Articles dans des conférences internationales avec comité de lecture et actes

Les deux premiers articles sont liés aux contributions présentées au Chapitre 2, tandis que le dernier est lié à celles présentées au Chapitre 3 :

- B. Larrousse and S. Lasaulce, *Coded Power Control : Performance Analysis*, IEEE Intl. Symposium on Information Theory (ISIT), Istanbul, Turkey, Jul. 2013.
- B. Larrousse, A. Agrawal, and S. Lasaulce, *Implicit coordination in 2-agent team problems. Application to distributed power allocation*, IEEE 12th Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), Hammamet, Tunisia, May 2014.
- B. Larrousse, O. Beaude, and S. Lasaulce, *Crawford-Sobel Meets Lloyd-Max on the Grid*, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, May 2014.

Chapitre 2

Coordination via la communication implicite dans les réseaux distribués.

« The more improbable the message, the less "compressible" it is, and the more bandwidth it requires. This is Shannon's point : the essence is its improbability. »

– William Poundstone, *Fortune's Formula : The Untold Story of the Scientific Betting System That Beat the Casinos And Wall Street (2006).*

Ce chapitre est consacré à la communication implicite, via les actions, entre plusieurs agents qui cherchent à se coordonner. Pourtant assez naturelle dans la littérature économique et utilisée plus ou moins implicitement dans la vie de tous les jours, comme expliqué au Chapitre 1, l'idée de communiquer via les actions ne semblent pas encore bien encré dans la littérature sur les communications sans fil, à part quelques exceptions près sur des exemples spécifiques de contrôle (voir par exemple [47]). Nous étudions ici un problème d'optimisation distribuée ayant une certaine structure d'information (qui sera interprétée comme une structure de communication implicite), et appliquons nos résultats à des problèmes de communication sans fil et notamment de contrôle de puissance.

Nota Bene : toutes les preuves des résultats de ce chapitre peuvent être trouvées dans l'article disponible dans l'Annexe E, section E.1.1.

2.1 Motivation et état de l'art

2.1.1 Un problème d'optimisation distribuée

Le problème principal étudié dans ce chapitre est un problème d'*optimisation distribuée*, dont nous détaillons le modèle en Section 2.2. La motivation pour ce problème est la suivante : considérons un entier $T \geq 1$, trois alphabets discrets $\mathcal{X}_0$, $\mathcal{X}_1$, $\mathcal{X}_2$ et une fonction d'utilité

$$w : \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \to \mathbb{R}. \tag{2.1}$$

Le but est de maximiser l'utilité moyenne, défini par l'équation suivante :

$$W_T(x_0^T, x_1^T, x_2^T) \triangleq \frac{1}{T}\sum_{t=1}^{T} w(x_{0,t}, x_{1,t}, x_{2,t}) \tag{2.2}$$

par rapport aux séquences $x_1^T \triangleq (x_{1,1}, ..., x_{1,T}) \in \mathcal{X}_1^T$ et $x_2^T \triangleq (x_{2,1}, ..., x_{2,T}) \in \mathcal{X}_2^T$ sachant la connaissance de la séquence $x_0^T \triangleq (x_{0,1}, ..., x_{0,T}) \in \mathcal{X}_0^T$. Si nous ne restreignons pas d'avantage le problème, avec la connaissance avant chaque étape t, $1 \leq t \leq T$, de $x_{0,t}$, résoudre ce problème d'optimisation consiste à trouver une des paires de variables $(x_{1,t}^\star, x_{2,t}^\star)$ pour chaque étape t qui induisent une utilité maximum. Cette configuration sera notre cas de **communication sans coût**, qui servira de référence par la suite. La valeur du maximum de W_T sera alors

$$W_T^\star = \frac{1}{T}\sum_{t=1}^{T} \max_{x_1, x_2} w(x_{0,t}, x_1, x_2). \tag{2.3}$$

Dans ce chapitre, nous ajoutons une contrainte au problème, qui est la suivante : la séquence x_2^T ne peut pas être contrôlée ou optimisée directement avec la séquence x_1^T. En effet, nous supposons qu'un agent (l'Agent 1) optimise la séquence x_1^T, qu'un autre agent (l'Agent 2) optimise la séquence x_2^T, et enfin que ces deux agents n'ont pas la même connaissance sur la séquence x_0^T : l'Agent 1 sera mieux informé que l'Agent 2. Nous supposons donc ici une **asymétrie d'information** entre les agents. L'Agent 1 va alors coder de l'information dans ces choix $x_{1,t}$, $1 \leq t \leq T$, que l'Agent 2 pourra décoder. Il est important de noter ici que les choix de l'Agent 1 ont une influence *directe* sur l'utilité moyenne W_T, ces choix devront donc être fait de façon à trouver le bon compromis entre transmission d'information et maximisation de l'utilité moyenne. Une des questions est alors de savoir comment coder et décoder l'information de manière optimale, c'est-à-dire pour maximiser l'utilité. C'est ici que la théorie de l'information intervient et en particulier les théorèmes de codage de source et de codage canal (Propositions A.13 et A.14). Cette restriction sur la connaissance (asymétrique) des agents induira une **contrainte d'information** dans le problème d'optimisation.
Un des résultats principaux de ce chapitre est de caractériser cette contrainte lorsque T tend vers l'infini et x_0^T est une séquence indépendante et identiquement distribuée (i.i.d.) de réalisations d'une variable aléatoire X_0. Cette dernière peut être vue comme notre variable d'*état de nature* que nous avons introduite au Chapitre 1. Cette variable est hors de contrôle des agents.

Ce problème est un cas particulier d'un problème d'optimisation distribuée, dans lequel K agents, $K \geq 2$, sont connectés via un graphe de communication et ont l'objectif commun de maximiser l'utilité moyenne W_T pour T grand. Dans ce cadre général, la variable x_k, avec $k \in \{1, \cdots, K\}$ est appelée l'*action* de l'Agent k et représente la seule variable qu'il peut contrôler. Le **graphe de communication** définit comment les agents interagissent en fonction de l'observation de l'état aléatoire et des actions des autres agents. L'utilité moyenne mesurera donc ici le degré de *coordination* entre les agents, sous les contraintes d'observations imposées par le graphe de communication. Une des questions centrales est donc de caractériser les valeurs possibles de cette utilité moyenne W_T quand les agents interagissent un grand nombre de fois, c'est-à-dire quand T est très grand.
Résoudre ce problème en toute généralité nous paraît très difficile, aussi ce chapitre se concentre sur le cas particulier où deux agents interagissent ($K = 2$). Nous considérons donc le cas où l'Agent 1 connaît parfaitement les réalisations passées, présentes, et futures de l'état x_0^T, et l'Agent 2 observe une version bruité (une observation imparfaite) des actions de l'Agent 1. De plus, nous considérons deux cas de connaissance sur l'état pour l'Agent 2 :

I. une connaissance strictement causale, c'est-à-dire des réalisations passées uniquement ;

II. aucune connaissance, même des réalisations passées.

Malgré ces hypothèses restrictives, nous pensons qu'il est possible d'extraire de ce travail des concepts et idées nouvelles qui auront un intérêt pratique, et qui pourront être appliqués au cas général avec plus de deux agents et avec un graphe de communication arbitraire.

2.1.2 État de l'art

Dans une grande partie de la littérature sur la coordination entre agents, qui inclue les problèmes de décision collective [13] [6], une des hypothèses est la possibilité pour les agents d'accéder à un **canal dédié** à la communication pour coordonner leurs actions. Ces canaux sont *dédiés* car ils permettent aux agents de communiquer entre eux sans affecter *directement* la fonction d'utilité. Les travaux sur la coordination avec canaux dédiés les plus proches des travaux de ce chapitre sont [29, 30]. Dans ces travaux, les auteurs introduisent les notions de *coordination empirique* et de *coordination forte* pour mesurer la capacité des agents à coordonner leurs actions dans un réseau composé de canaux dédiés sans bruit. La coordination empirique mesure un comportement moyen de coordination au cours du temps et impose à la distribution empirique conjointe des actions d'être proche asymptotiquement en distance de variation totale (voir la Définition A.1 disponible en Annexe A) d'une distribution cible. Cette notion est liée à des travaux précédents sur la communication de distribution de probabilité [57], et est analysée via des outils de théorie taux-distorsion [14]. La coordination forte est quant à elle plus contraignante et impose que les distributions des *séquences* d'actions soient asymptotiquement indiscernables (également en distance de variation totale) des séquences d'actions tirées grâce à une distribution cible. Cette notion est quant à elle liée au concept de « channel resovability » [49]. Dans les deux cas, le but est d'établir la *capacité de coordination*, qui lie les distributions conjointes d'actions atteignables aux contraintes sur le taux fixé des canaux dédiés non bruités. Les résultats de [29, 30] ont été étendus à un certain nombre de réseaux avec canaux dédiés [11, 13, 14, 48], et des codes optimaux ont été conçus pour des configurations particulières [12, 15].

En revanche, beaucoup moins de choses sont connues sur le problème de coordination via les actions (la communication implicite) en l'absence de canaux dédiés à la communication. Le travail le plus proche est celui de [46], dans lequel les auteurs caractérisent l'ensemble des utilités moyennes possibles, en supposant que l'Agent 2 observe parfaitement les actions passées de l'Agent 1. Ils établissent l'ensemble des *distributions atteignables*, qui sont en fait les distributions empiriques conjointes des actions que l'on peut atteindre via la structure d'information imposée au système. En particulier, cet ensemble est caractérisé par une contrainte d'information qui traduit la structure d'observation des actions entre les agents. Cet article [46] s'appuie en grande partie sur des arguments de type combinatoire, alors que [32] utilise une approche via des outils de la théorie de l'information pour la coordination via les actions. Nous utiliserons nous aussi des outils issus de la théorie de l'information.

Concernant l'application développée dans ce chapitre, ce travail apparaît comme le premier à établir une relation entre le problème d'optimisation distribuée étudié ici et le problème de contrôle/d'allocation distribué(e) de ressources dans les réseaux sans fil. Spécifiquement, nous

13. « Team decision problems » en anglais.
14. « Rate-distortion theory » en anglais

nous focalisons sur un problème de contrôle de puissance distribué pour un canal avec interférence, ainsi qu'un cas particulier de ce canal, le canal à accès multiple (voir la définition en Section A.5). Dans la littérature, un point de vue souvent adopté est d'utiliser des outils de théorie des jeux pour concevoir des schémas de contrôle de puissance et pour analyser leur performance. Un de ces outils est l'algorithme itératif de « water-filling » [87], qui est un cas particulier de la dynamique de meilleure réponse (un algorithme de la théorie des jeux, voir la section B.3 en Annexe B), qui est appliqué sur un horizon de temps sur lequel l'état du canal est constant.
Un des principaux inconvénients des implémentations diverses de la dynamique de meilleure réponse pour des problèmes de contrôle de puissance (voir par exemple, [58][88][5]) sont qu'ils induisent (après convergence) des stratégies de contrôle de puissance amenant à des équilibres de Nash non efficace au sens de Pareto. C'est-à-dire qu'il existe des stratégies qui permettraient à tous les agents d'augmenter leur utilité individuelle par rapport à la stratégie de « contrôle de puissance Nash ». Un des autres inconvénients est que ce genre de stratégies itératives ne convergent pas toujours. Il existe seulement des conditions suffisantes de convergence (voir par exemple [78] pour le cas « entrées multiples, sorties multiples » [15]), et ces conditions peuvent être très restrictives à tel point qu'elles peuvent ne pas être vérifiées du tout dans certains cas importants tels que le canal à accès multiple parallèle [69].
D'autre part, un des buts du **contrôle de puissance codé**, qui est le nouveau concept développé dans ce chapitre (et en particulier en Section 2.5), et d'obtenir un point de fonctionnement efficace pour le réseau. Ceci est rendu possible en échangeant de l'information entre les émetteurs à propos de l'état du canal. Une observation clé que nous faisons est que l'échange d'information peut être fait via des quantités observées telles que le rapport signal sur interférence plus bruit (RSIB) [16]. C'est-à-dire que ces rapports, pour les différents utilisateurs, peuvent être eux-mêmes vus comme les sorties d'un canal par lequel les émetteurs peuvent communiquer pour coordonner leurs actions, à condition qu'ils encodent ces rapports de manière appropriée, ce qui n'est pas fait dans la littérature actuelle. L'aspect de codage apparaît ici car un émetteur va prendre plusieurs réalisations de l'état du canal pour les mettre en correspondance avec une séquence de niveaux de puissance. Les autres émetteurs exploitent alors la séquence d'observations associée (par exemple les valeurs de leur RSIB) pour sélectionner leurs niveaux de puissance. L'avantage d'utiliser du codage est qu'aucune procédure itérative n'est nécessaire (les livres de codes et les fonctions de codage et de décodage se font hors ligne, avant l'utilisation du réseau), nous évitant donc des problèmes de convergence.
L'efficacité du contrôle de puissance codé sera étudiée dans ce chapitre, et sera notamment comparé aux performances du contrôle de puissance Nash. D'autres aspects tels que la disponibilité des informations sur l'état du canal et la complexité devraient être considérés, mais ceci est laissé comme une extension complète des travaux de ce chapitre. Notamment car ces aspects requièrent une structure d'information *symétrique*, ce qui n'est pas le cas dans ce chapitre.

2.1.3 Contributions principales

Les principales contributions de ce chapitre sont les suivantes :

- Premièrement, les résultats de la Section 2.3 étendent ceux de [46] en relâchant certaines hypothèses faites sur le graphe de communication. Spécifiquement, alors que les auteurs de [46] ont besoin dans leur définition des stratégies de la *mémoire* sur les séquences d'actions passées des deux agents, nous relâchons cette hypothèse. Nous supposons que l'Agent 1 n'a

15. « Multiple-Input Multiple-Output » (MIMO) en anglais.
16. « Signal-to-noise plus interference ratio » (SINR) en anglais.

aucune mémoire sur ses propres actions passées ainsi que sur celles de l'Agent 2, et que l'Agent 2 n'a aucune mémoire sur ses propres actions passées, mais une certaine connaissance sur les actions passées de l'Agent 1. En effet, alors que [46] suppose que l'Agent 2 observe parfaitement les actions passées de l'Agent 1, nous considérons le cas où il observe de manière imparfaite ces actions. Seule une version bruitée de ces actions est observée. Enfin, comme expliqué précédemment, nous étudions une situation similaire à celle considérée par les auteurs de [46] pour laquelle l'Agent 2 a une connaissance strictement causale des réalisations de l'état (Théorème 2.7), ainsi qu'une situation plus générale où l'Agent 2 n'a aucune connaissance des réalisations de l'état (Théorème 2.8).

- Deuxièmement, nous clarifions les liens entre la formulation type théorie des jeux de [46] et une formulation de type théorie de l'information. En particulier, nous discutons des liens entre les théorèmes de codage de la littérature sur les canaux dépendants d'un état [42, 55, 20, 21, 22], les théorèmes de séparations, et les articles traitant de coordination empirique [30, 31]. De plus, le problème de la détermination de la valeur de l'utilité moyenne de long terme est formulé comme un problème d'optimisation qui est expliqué en détail dans la Section 2.4. Cette étude, qui prouve notamment que le problème est convexe, nous permet d'effectuer une étude numérique approfondie (en Section 2.5) pour le problème de contrôle de puissance distribué, montrant ainsi les bénéfices potentiels de l'approche proposée.

- Troisièmement, nous établissons une connexion entre le problème de coordination via les actions et le problème de contrôle de puissance dans les réseaux sans fil. Nous développons une nouvelle perspective pour l'allocation/le contrôle de ressources, dans laquelle concevoir une stratégie d'allocation de ressources induisant de bonnes performances en terme de l'utilité moyenne (commun à tous les agents du système) revient à concevoir un code. Ce code sera bon (en terme d'utilité moyenne) s'il implémente un bon compromis entre transmission d'information (à propos des réalisations futures de l'état, qui sera donc la *source d'information*, et qui permettra d'atteindre de bonnes valeurs de l'utilité future) et maximisation de l'utilité courante, c'est-à-dire le paiement d'étape. La Section 2.5.5 fournit en particulier une description complète d'un code de contrôle de puissance pour le canal à accès multiple.

2.2 Description du problème étudié

2.2.1 Formulation mathématique

Dans cette section, nous introduisons formellement le problème étudié dans le reste de ce chapitre. Nous considérons deux agents qui doivent sélectionner leurs actions de manière répétée sur $T \geq 1$ étapes ou intervalles de temps et veulent se coordonner via leurs actions tout en prenant en compte la présence d'un état aléatoire qui a une influence sur l'utilité moyenne des agents. La coordination se fait via un graphe de communication que nous détaillons ci-après. A chaque étape $t \in \{1, \cdots, T\}$, l'action de l'Agent $k \in \{1, 2\}$ est $x_{k,t} \in \mathcal{X}_k$ avec $|\mathcal{X}_k| < \infty$, et la réalisation de l'état aléatoire est $x_{0,t} \in \mathcal{X}_0$ avec $|\mathcal{X}_0| < \infty$. Nous ne considérons donc ici que des ensembles *finis*. Les réalisations de l'état sont supposées être i.i.d. selon une variable aléatoire X_0 de distribution de probabilité ρ_0. L'état aléatoire ne dépend pas des actions des agents mais

affecte la fonction [17] (commune aux deux agents) de l'utilité moyenne

$$w : \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \to \mathbb{R}. \tag{2.4}$$

La coordination est mesurée en terme de l'utilité moyenne $W_T(x_0^T, x_1^T, x_2^T)$ comme défini à l'équation (2.2).
Nous supposons donc qu'à chaque étape t, l'Agent 2 a seulement accès à des observations *imparfaites* $y_t \in \mathcal{Y}$, avec $|\mathcal{Y}| < \infty$, des actions de l'Agent 1. Spécifiquement, la séquence d'observations y^T est la sortie d'un canal sans mémoire ayant pour probabilité de transition

$$\mathrm{P}(y^T | x_0^T, x_1^T, x_2^T) = \prod_{t=1}^{T} \Gamma(y_t | x_{0,t}, x_{1,t}, x_{2,t}) \tag{2.5}$$

pour une distribution de probabilité conditionnelle Γ quelconque fixée.
Nous considérons deux **graphes de communication asymétriques** qui restreignent l'observation de l'état par les agents ainsi que les observations pour chaque agent des actions de l'autre agent. Ces graphes sont caractérisés par les stratégies $(\sigma_t)_{1\leq t\leq T}$ et $(\tau_t)_{1\leq t\leq T}$ de l'Agent 1 et de l'Agent 2 respectivement. Ces stratégies sont des séquences d'applications définies pour toutes étapes $t \in \{1, ..., T\}$ par :

$$\text{Cas I} : \begin{cases} \sigma_t^{\mathrm{I}} & : & \mathcal{X}_0^T & \to & \mathcal{X}_1 \\ \tau_t^{\mathrm{I}} & : & \mathcal{X}_0^{t-1} \times \mathcal{Y}^{t-1} & \to & \mathcal{X}_2 \end{cases} \tag{2.6}$$

$$\text{Cas II} : \begin{cases} \sigma_t^{\mathrm{II}} & : & \mathcal{X}_0^T & \to & \mathcal{X}_1 \\ \tau_t^{\mathrm{II}} & : & \mathcal{Y}^{t-1} & \to & \mathcal{X}_2 \end{cases} . \tag{2.7}$$

où le Cas I correspond à l'observation strictement causale des réalisations de l'état par l'Agent 2, alors que le cas II correspond au cas sans observations de ces réalisations. Comme nous le voyons, dans les deux cas, les séquences d'actions passées des deux agents ne sont pas en entrée des applications. Seule une version bruitée des actions passées de l'Agent 1 est en entrée des stratégies de l'Agent 2 (et uniquement de l'Agent 2).
Notre objectif est de caractériser l'ensemble des utilités moyennes qui sont réalisables asymptotiquement, c'est-à-dire les valeurs possibles pour

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} w(x_{0,t}, x_{1,t}, x_{2,t}) \tag{2.8}$$

sous les contraintes imposées par les deux graphes de communication, respectivement définies par les équations (2.6) et (2.7). La définition de ces deux ensembles de paiements réalisables est la suivante :

Définition 2.1 (Ensembles des paiements réalisables). *L'ensemble des paiements réalisables dans le Cas I est défini par*

$$\Omega^{\mathrm{I}} = \left\{ \omega \in \mathbb{R} \ : \ \exists (\sigma_t^{\mathrm{I}}, \tau_t^{\mathrm{I}})_{1\leq t\leq T}, \ \omega = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} w\left(x_{0,t}, \sigma_t^{\mathrm{I}}(x_0^T), \tau_t^{\mathrm{I}}(y^{t-1}, x_0^{t-1})\right) \right\}. \tag{2.9}$$

L'ensemble des paiements réalisables dans le Cas II est défini par

$$\Omega^{\mathrm{II}} = \left\{ \omega \in \mathbb{R} \ : \ \exists (\sigma_t^{\mathrm{II}}, \tau_t^{\mathrm{II}})_{1\leq t\leq T}, \ \omega = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} w\left(x_{0,t}, \sigma_t^{\mathrm{II}}(x_0^T), \tau_t^{\mathrm{II}}(y^{t-1})\right) \right\}. \tag{2.10}$$

17. La fonction w peut être n'importe quelle fonction telle que les utilités moyennes asymptotiques définis dans ce chapitre existent.

Les ensembles des paiements réalisables sont directement liés aux ensembles de distributions empiriques conjointes réalisables, distributions sur l'ensemble :

$$\mathcal{X} \triangleq \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2, \tag{2.11}$$

Ces ensembles de distributions empiriques conjointes réalisables sont définis par

Définition 2.2 (Atteignabilité). *Pour $\ell \in \{\mathrm{I}, \mathrm{II}\}$, la distribution de probabilité $\overline{Q}(x_0, x_1, x_2)$ est atteignable s'il existe une paire de stratégies $(\sigma_t^\ell, \tau_t^\ell)_{1 \leq t \leq T}$ induisant à chaque étape t une distribution conjointe empirique*

$$\mathrm{P}_{X_{0,t}, X_{1,t}, X_{2,t}, Y_t}(x_0, x_1, x_2, y) \triangleq \Gamma(y|x_0, x_1, x_2)\mathrm{P}_{X_{1,t}, X_{2,t}|X_{0,t}}(x_1, x_2|x_0)\rho_0(x_0), \tag{2.12}$$

telle que pour tout $(x_0, x_1, x_2) \in \mathcal{X}$

$$\lim_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{y \in \mathcal{Y}} \mathrm{P}_{X_{0,t}, X_{1,t}, X_{2,t}, Y_t}(x_0, x_1, x_2, y) = \overline{Q}(x_0, x_1, x_2). \tag{2.13}$$

c'est-à-dire que la distribution conjointe empirique converge vers $\overline{Q}$ avec T.

Chaque ensemble des paiements est **l'image linéaire** de l'ensemble de distributions atteignables correspondant via l'opérateur espérance $\mathbb{E}$. Par conséquent, une valeur quelconque fixée ω est réalisable asymptotiquement si et seulement s'il existe une distribution atteignable $\overline{Q}$ telle que $\omega = \mathbb{E}_{\overline{Q}}[w]$. Dans ce chapitre, nous nous concentrerons donc sur la caractérisation des distributions atteignables plutôt que sur une caractérisation directe des ensembles des paiements réalisables.

2.2.2 Discussion

La notion de distribution atteignable peut aussi être connectée à la notion de distribution empirique atteignable [30], via le **type** (voir Définition A.16 en annexe) des séquences d'actions. En dénotant T_{X^T} la distribution obtenue en construisant l'*histogramme* de la séquence x^T, il s'ensuit grâce à la Définition 2.2 d'atteignabilité qu'une distribution de probabilité $\overline{Q}$ sur $\mathcal{X} = \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2$ est atteignable s'il existe une séquence de stratégies $(\sigma_t^\ell, \tau_t^\ell)_{1 \leq t \leq T}$, $\ell \in \{\mathrm{I}, \mathrm{II}\}$ qui induit, avec la séquence x_0^T, la séquence $x^T \in \mathcal{X}$ telle que

$$\lim_{T \to \infty} ||\mathbb{E}(\mathrm{T}_{X^T}) - \overline{Q}||_1 = 0, \tag{2.14}$$

c'est-à-dire que l'histogramme moyen d'une séquence d'actions et de réalisations de l'état est arbitrairement proche de la distribution de probabilité $\overline{Q}$.

Rappelons la définition de **coordination empirique** pour les cas I et II :

Définition 2.3 (Distribution empirique [30]). *Pour $\ell \in \{\mathrm{I}, \mathrm{II}\}$, une distribution de probabilité $\overline{Q}$ sur $\mathcal{X}$ est une coordination empirique atteignable s'il existe une séquence de stratégies $(\sigma_t^\ell, \tau_t^\ell)_{1 \leq t \leq T}$ qui induit, avec la séquence x_0^T, la séquence $x^T \in \mathcal{X}$ telle que*

$$\forall \epsilon > 0, \quad \lim_{T \to \infty} \mathrm{P}(||\mathrm{T}_{X^T} - \overline{Q}||_1 > \epsilon) = 0, \tag{2.15}$$

c'est-à-dire que la distance entre l'histogramme de la séquence d'actions et $\overline{Q}$ converge en probabilité vers 0.

La proposition suivante fait le lien entre les notions de distribution atteignable et coordination empirique atteignable :

Proposition 2.4. *Si $\overline{Q} \in \Delta(\mathcal{X})$ est une coordination empirique atteignable, alors c'est une distribution atteignable.*

Avant d'analyser le problème de coordination en détail dans la Section 2.3, quelques commentaires sont nécessaires.

- Dans [46], les auteurs supposent que les stratégies sont données par deux séquences d'applications

$$\begin{array}{rlll} & \sigma_t : & \mathcal{X}_0^T \times \mathcal{X}_1^{t-1} \times \mathcal{X}_2^{t-1} & \to \mathcal{X}_1 \\ \text{et} & \tau_t : & \mathcal{X}_0^{t-1} \times \mathcal{X}_1^{t-1} \times \mathcal{X}_2^{t-1} & \to \mathcal{X}_2 \end{array} \tag{2.16}$$

  et que l'Agent 2 observe parfaitement la séquence d'actions de l'Agent 1, c'est-à-dire $Y = X_1$. Un de nos résultats est tel que, sous l'hypothèse d'observation parfaite, il n'y a pas de perte de généralité à considérer les stratégies données par (2.7) au lieu de (2.16). Cependant, l'ensemble des équilibres de Nash atteignables obtenu avec (2.7) peut ne pas coïncider avec celui associé aux stratégies (2.16) de [46]. Caractériser l'ensemble des équilibres de Nash atteignables est pertinent lorsque les agents ont des intérêts divergents. Lorsque c'est le cas, il s'avère que le fait qu'un agent puisse observer les actions passées des autres agents a une importance. Se restreindre aux stratégies (2.6) ou (2.7) nous fera donc perdre en généralité si les intérêts des agents sont divergents. Par exemple, dans une configuration de contrôle de puissance, si les émetteurs implémentent un schéma de coopération qui fait transmettre à faible puissance tant qu'aucun émetteur n'utilise un niveau de puissance élevé (voir par exemple [83]), il est important pour les émetteurs d'être capable de vérifier que les autres émetteurs utilisent effectivement ce schéma de coopération et émettent à faible niveau de puissance.

- Il est aussi important de noter que les stratégies ne sont pas des stratégies par bloc comme c'est conventionnellement le cas pour des stratégies de codage et de décodage. Ici, un agent doit agir à toutes les étapes (comme dans un jeu répété), et les stratégies considérées peuvent être vues comme des stratégies de codage source-canal conjoint avec un codage/décodage à la volée.

- Enfin, nous soulignons que la stratégie de l'Agent 1 suppose une connaissance non causale des réalisations de l'état. Ce genre d'hypothèse a été utilisée en théorie de l'information, par exemple par Gel'fand et Pinsker [42] pour le codage à travers un canal dépendant d'un état. Leur codage est complètement pertinent pour le problème de codage à travers des canaux de diffusion [18][40] et pour des problèmes de filigrane[18] [23]. Une des principales contributions développée dans ce chapitre en Section 2.5 étant dédiée à l'application de ce modèle de communication à des problèmes de contrôle de puissance, nous ajoutons plusieurs motivations pour justifier l'hypothèse de non causalité sur les réalisations de l'état. Premièrement, même si l'Agent 1 ne connaît les réalisations futures que sur un horizon de temps réduit (connaître une réalisation en avance est déjà très utile dans les problèmes de contrôle de puissance), le réseau peut être coordonné à un niveau qui induit de bonnes performances par rapport aux approches conventionnelles telles que l'utilisation d'une stratégie distribuée qui implémente un équilibre de Nash (voir par exemple [44][87][78][58] pour

18. « Watermarking problem » en anglais

des exemples de contrôle de puissance distribué). La situation la plus commune en contrôle de puissance est de supposer que deux phases sont disponibles, une phrase d'apprentissage et une phase d'action, et qu'une réalisation d'état est connue à l'avance pour ajuster le niveau de puissance. Ce cas particulier peut être obtenu en prenant $T = 2$, c'est-à-dire $t \in \{1, 2\}$. Pour ce type de scénario, comme illustré en Figure 2.1, une stratégie simple de coordination lorsque l'Agent 1 connaît la réalisation suivante de l'état peut être d'informer l'Agent 2 sur cette réalisation [19] aux étapes impaires et de coordonner leurs actions aux étapes paires.

Le fait que l'Agent 1 connaisse toutes les réalisations de l'état en avance peut aussi être vu comme un manière d'obtenir une borne supérieure sur la performance de scénarios dans lesquels l'horizon de temps pour les prédictions de l'état est réduit. Il est important de noter que des scénarios pratiques sont apparus ces dernières années, pour lesquels les prédictions de l'état sur un horizon de temps long est une hypothèse réaliste. Par exemple, cela devient de plus en plus courant d'exploiter la trajectoire prédite d'un utilisateur mobile pour optimiser un système [39], ce qui rend notre approche pertinente si nous interprétons l'état de canal comme l'atténuation.

Enfin, l'approche que nous proposons peut aussi être appliquée au cas où l'état n'est pas i.i.d. d'étape en étape mais i.i.d. de bloc en bloc, un bloc étant constitué de plusieurs étapes. En effet, il existe des standards de communication sans fil qui supposent que l'état du canal est constant sur plusieurs étapes, et notre approche suggère que des gains peuvent être obtenus en faisant varier le niveau de puissance d'étape en étape même si l'état du canal est constant.

19. Par exemple, l'Agent 1 peut utilisé un niveau de puissance élevé (respectivement faible) sur une étape impaire pour informer l'Agent 2 que le canal est bon (respectivement mauvais) à l'étape paire suivante.

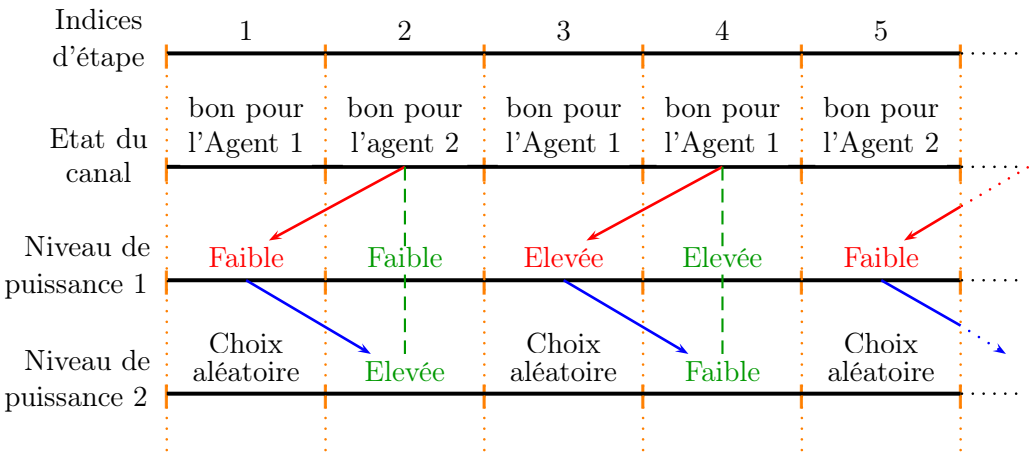

FIGURE 2.1 – **Schéma de coordination simple** : ce schéma illustre une stratégie simple de coordination entre deux émetteurs (les agents) pour un scénario simplifié avec alphabets binaires $\mathcal{X}_0 = \{\text{bon pour l'Agent 1 }, \text{bon pour l'Agent 2 }\}$, $\mathcal{X}_1 = \mathcal{X}_2 = \{\text{Faible}, \text{Elevée}\}$. L'émetteur informé (l'Agent 1) choisit la puissance de transmission la plus faible (respectivement la plus élevée) sur l'étape courante *impaire* $2t'+1$ si la réalisation de l'état du canal de l'étape suivante $2t'+2$ est bonne pour l'Agent 2 (respectivement l'Agent 1). Si l'Agent 2 peut retrouver parfaitement les niveaux de puissance de l'Agent 1, il connaît donc la réalisation de l'état sur les étapes paires. Il transmet alors à puissance faible (respectivement élevée) si le canal est bon pour l'Agent 1 (respectivement 2). Pour les étapes impaires, il choisit une puissance d'émission aléatoirement. Voir aussi la Figure 1.8.

2.3 Contraintes d'information sur les distributions atteignables

Dans cette section, nous caractérisons les ensembles de distributions conjointes empiriques atteignables $\overline{Q} \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ pour les cas décrits par les équations (2.6) et (2.7). Nous montrons que ces ensembles sont constitués de distributions de probabilité sur $\Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ qui vérifient une **contrainte d'information** dont nous donnons une caractérisation explicite et qui capture les restrictions imposées par le graphe de communication. Nous commençons dans la Section 2.3.1 par fournir une condition *nécessaire* d'atteignabilité (Théorème 2.5), valable pour les cas I et II, puis en Section 2.3.2 nous explicitons une condition *suffisante* d'atteignabilité pour le cas I (Théorème 2.7) ainsi que pour le cas II (Théorème 2.8).

2.3.1 Une condition nécessaire d'atteignabilité

Le théorème suivant explicite donc une condition nécessaire d'atteignabilité, aussi bien pour le cas I que pour le cas II.

Théorème 2.5. *Soit $\overline{Q}$ une distribution de probabilité sur $\Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ telle que*

$$\sum_{x_1,x_2} \overline{Q}(x_0,x_1,x_2) = \rho_0(x_0) \; \forall \, x_0 \in \mathcal{X}_0. \tag{2.17}$$

Aussi bien dans le cas I que dans le cas II, une distribution de probabilité $\overline{Q}$ est atteignable si

c'est la distribution marginale de $Q \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})$ *qui se factorise comme suit :*

$$Q(x_0, x_1, x_2, y) = \Gamma(y|x_0, x_1, x_2)\overline{Q}(x_0, x_1, x_2),\ \forall\,(x_0, x_1, x_2, y) \in \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y} \quad (2.18)$$

et qui vérifie la contrainte d'information :

$$I_Q(X_0; X_2) \leq I_Q(X_1; Y|X_0, X_2) \quad (2.19)$$

*où les définitions d'***information mutuelle** *et d'***information mutuelle conditionnelle** *sont données en Annexe A.1 aux Définitions A.7 et A.10 respectivement.*

La démonstration de ce théorème n'est pas détaillée ici. Elle utilise des outils de théorie de l'information tels que des règles de chaîne (Définitions A.6 et A.11) ainsi que les propriétés de l'entropie (Proposition A.9). De plus, l'hypothèse i.i.d. sur l'état ainsi que la définition de la stratégie de l'Agent 2 (dans le cas I aussi bien que dans le cas II) sont cruciales dans cette démonstration. Pour conclure cette dernière, nous avons également besoin du Lemme de convexité 2.6 qui suit. Avant de l'énoncer, introduisons la fonction Φ^{I} :

$$\Phi^{\mathrm{I}} : \left| \begin{array}{ccc} \Delta(\mathcal{X} \times \mathcal{Y}) & \to & \mathbb{R} \\ Q & \mapsto & I_Q(X_0; X_2) - I_Q(X_1; Y|X_0, X_2) \end{array} \right. . \quad (2.20)$$

Nous avons alors :

Lemme 2.6. *La fonction* Φ^{I} *est strictement convexe sur l'ensemble des distributions de probabilité* $Q \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})$ *qui vérifient*

$$\sum_{x_1, x_2, y} Q(x_0, x_1, x_2, y) = \rho_0(x_0),\ \forall\, x_0 \in \mathcal{X}_0 \quad (2.21)$$

et

$$Q(x_0, x_1, x_2, y) = \Gamma(y|x_0, x_1, x_2)\mathrm{P}(x_0, x_1, x_2),\ \forall\,(x_0, x_1, x_2, y) \in (\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}) \quad (2.22)$$

avec les distributions ρ_0 *et* Γ *quelconques fixées.*

Ce lemme, dont la démonstration peut être trouvée dans l'article de la section E.1.1, sera également utile pour caractériser le problème d'optimisation étudié en Section 2.4.

Le théorème 2.5 peut être interprété comme suit. Les actions de l'Agent 2, représentées par X_2, correspondent à la séquence décodée d'un problème de source-canal conjoint avec distorsion sur la source d'information, la source étant représentée par X_0. Le canal a pour entrées les actions de l'Agent 1, représentée par X_1, et pour sorties le signal Y observé par l'Agent 2. Pour être atteignable, le taux de distorsion doit être plus petit que le taux de transmission du canal, comme il est classique en théorie de l'information (voir le théorème A.15). Dès lors, la paire de variables $S = (X_0, X_2)$ peut être vue comme une **information adjacente**, comme nous pouvons trouver dans le problème de canal avec dépendance à un état [40]. Comme nous le verrons en Section 2.3.2, nous utilisons cette interprétation pour établir les conditions suffisantes d'atteignabilité. En revanche, pour prouver notre condition nécessaire, cet argument ne peut être utilisé. En effet, un des arguments classiques pour les preuves de réciproques (les conditions nécessaires donc) pour les canaux avec dépendance à un état [42, 66] est d'utiliser l'*indépendance* entre le « message » transmis et l'état. Or ici le rôle du message est joué par la quantité X_0^T, qui n'est clairement pas indépendant de $S^T = (X_0^T, X_2^T)$, que nous voudrions interpréter comme l'information adjacente. C'est une des raisons pour lesquelles la preuve de la réciproque ne suit pas de résultats existants et a dû être repensée.

La section suivante énonce les conditions suffisantes d'atteignabilité pour les cas I et II.

2.3.2 Conditions suffisantes d'atteignabilité

Avant de présenter les résultats sur les conditions suffisantes d'atteignabilité, un mot sur l'application considérée en Section 2.5 : la distribution de probabilité conditionnelle Γ représente la structure d'observation de l'Agent 2, c'est-à-dire qu'elle définit ce que l'Agent 2 observe sur la séquence d'actions de l'Agent 1. En appliquant ce cadre à un réseau à interférences avec deux émetteurs et deux récepteurs, l'émetteur 1 peut être vu comme l'agent informé (par exemple un émetteur principal [50] ou maitre [62]), et Γ peut représenter un canal avec retour RSIB [20] du récepteur 2 vers l'émetteur 2.

Atteignabilité : Cas I

Nous énonçons donc maintenant les théorèmes d'atteignabilité, à commencer par le cas I, c'est-à-dire les stratégies (2.6) pour les deux agents :

Théorème 2.7. *Considérons la structure d'observation du Cas I. Soit* $\overline{Q} \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ *avec* $\sum_{(x_1,x_2)} \overline{Q}(x_0,x_1,x_2) = \rho_0(x_0)\ \forall x_0 \in \mathcal{X}_0$. *Si la distribution de probabilité* Q, *structurée comme suit :*

$$Q(x_0,x_1,x_2,y) = \Gamma(y|x_0,x_1,x_2)\overline{Q}(x_0,x_1,x_2),\ \forall\ (x_0,x_1,x_2,y) \in \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y} \quad (2.23)$$

vérifie la contrainte :

$$I_Q(X_0;X_2) \leq I_Q(X_1;Y|X_0,X_2), \quad (2.24)$$

alors $\overline{Q}$ *est atteignable au sens de la Définition 2.2, pour des stratégies de la forme (2.6).*

La preuve d'atteignabilité exhibe des stratégies

$$f : \underline{x}_0^{(b+1)}, \underline{x}_0^{(b)}, \underline{x}_2^{(b)} \mapsto \underline{x}_1^{(b)} \quad (2.25)$$

et

$$g : \underline{x}_0^{(b)}, \underline{y}^{(b)}, \underline{x}_2^{(b)} \mapsto \underline{x}_2^{(b+1)} \quad (2.26)$$

de type bloc-Markov qui induisent une distribution atteignable lorsque la contrainte d'information (2.24) est vérifiée. A l'inverse de la démonstration réciproque du Théorème 2.5, trouver un code particulier est suffisant. Or, nous exhibons un code par bloc avec des mots de code suffisamment longs, ce qui permet de réutiliser les techniques standards de preuve de codage source avec distorsion et de codage canal. En effet, grâce au codage par bloc, ce qui joue le rôle du message à transmettre est la séquence d'états $\underline{x}_0^{(b+1)}$ du bloc *suivant* $b+1$, alors que la variable correspondant à l'information adjacente est $(\underline{x}_0^{(b)}, \underline{x}_2^{(b)})$, qui sont les séquences d'états et d'actions de l'Agent 2 du bloc *courant*. Nous pouvons donc bien utiliser ici l'argument d'indépendance entre le message source et l'information adjacente, et réutiliser des schémas classiques de codage. La méthodologie pour cette preuve est la suivante :

20. « SINR feedback channel » en anglais.

- nous supposons que le nombre d'étapes T est découpé en B blocs de taille n : $T = nB$ [21]. Avant chaque bloc b, $1 \leq b \leq B$, l'Agent 1 choisit sa séquence d'actions $\underline{x}_1^{(b)}$ en fonction de l'information à transmettre $\underline{x}_0^{(b+1)}$ (la séquence d'états du bloc suivant) et des informations adjacentes $(\underline{x}_0^{(b)}, \underline{x}_2^{(b)})$ dont il dispose sur le bloc courant. A la fin de chaque bloc b, l'Agent 2 observera lui $\underline{y}^{(b)}$ et choisira sa séquence d'actions pour le bloc suivant $\underline{x}_2^{(b+1)}$ en fonction de ses observations ainsi que de l'information adjacente du bloc courant $(\underline{x}_0^{(b)}, \underline{x}_2^{(b)})$.

- Les livres de codes ainsi que les séquences d'actions des agents sont générés grâce à la méthode des types [28], qui caractérisent notamment le nombre de séquences qui ont la même distribution empirique. Choisir ce type de séquences permet de créer un code dont la probabilité d'erreurs pourra être bornée par un nombre aussi proche de 0 que l'on veut, et ce grâce un certain nombre de lemmes qui sont présentés en Annexe A.4. Ces lemmes donnent des conditions sur le nombre de séquences qu'il faut considérer pour le livre de codes source ainsi que le livre de codes canal, ce qui se traduit finalement par la contrainte d'information (2.24). Cette méthode est centrale en théorie de l'information.

- Dans cette preuve, nous séparons les opérations de codage/décodage source et canal. Nous faisons donc tout d'abord le codage source, puis le codage canal, puis le décodage canal et enfin le décodage source, comme illustré en Figure 2.2. Le codage n'est donc pas un codage conjoint. Le fait que ces opérations soient séparées peut engendrer, dans certains systèmes comme celui où deux sources discrètes sans mémoire sont envoyées sur un canal à accès multiple discret et sans mémoire, une perte en terme de performance. Lorsque c'est le cas, il faut utiliser des codages conjoints, souvent plus difficiles à mettre en place.

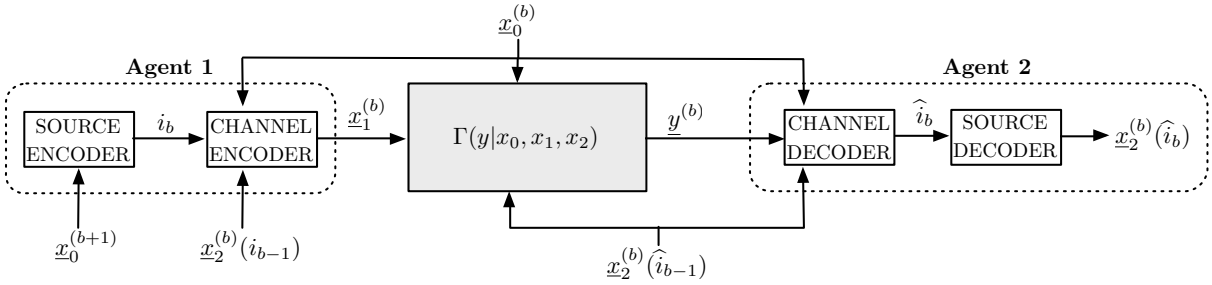

FIGURE 2.2 – **Illustration du schéma de codage** : cette figure illustre l'encodage source, l'encodage canal, le décodage canal, ainsi que le décodage source, au bloc b du schéma de codage par bloc proposé pour la démonstration du théorème 2.7.

La preuve d'atteignabilité montre que l'Agent 1, pour corréler ses actions sur le bloc b, utilise $(\underline{x}_0^{(b)}, \underline{x}_2^{(b)}(i_b))$ comme « information adjacente » , avec i_b l'indice correspondant à la séquence à transmettre $\underline{x}_0^{(b+1)}$. En revanche, l'Agent 2 utilise $(\underline{x}_0^{(b)}, \underline{x}_2^{(b)}(\widehat{i}_b))$, avec $\widehat{i}_b$ l'indice *décodé* par l'Agent 2 à la fin du bloc b, et qui a une probabilité positive d'être différent de i_b. Néanmoins, les agents ont la même information adjacente *asymptotiquement* sous la condition que Q vérifie la contrainte d'information (2.24). C'est à dire que asymptotiquement, $\widehat{i}_b = i_b$ avec une probabilité qui tend vers 1. C'est une observation clé dans la démonstration d'atteignabilité. Cela explique notamment pourquoi le fait que l'Agent 1 n'observe pas les actions de l'Agent 2 n'induit aucune

21. Nous supposons n et B entiers sans perte de généralité.

perte de performance.

Comme nous le voyons, la contrainte d'information (2.19) du Théorème 2.5 coïncide avec celle (2.24) du Théorème 2.7, ce qui veut dire que pour des stratégies correspondant au cas I, cette contrainte d'information est une condition nécessaire et suffisante pour qu'une distribution de probabilité conjointe $\overline{Q}(x_0, x_1, x_2)$ soit atteignable. Cela montre que séparer les opérations de codage/décodage source et canal n'implique aucune perte d'optimalité dans la preuve d'atteignabilité. Les Théorèmes 2.5 et 2.7 fournissent donc une caractérisation complète des distributions atteignables et donc par extension de ce qu'il est possible d'atteindre en terme d'utilité espérée (commun aux deux agents) lorsque la structure d'information est définie via les stratégies (2.6) et lorsque l'état aléatoire est i.i.d..
Lorsque l'état n'est pas i.i.d. d'étape en étape, mais constant sur $L \geq 1$ étapes consécutives et i.i.d. de bloc en bloc (pour des blocs contenant L étapes), une nouvelle contrainte d'information peut être obtenue en multipliant la partie gauche de (2.19) par $\frac{1}{L}$:

$$\frac{1}{L} I_Q(X_0; X_2) < I_Q(X_1; Y | X_0, X_2) \tag{2.27}$$

Cette contrainte modifiée est utile pour les communications sans fil par exemple, lorsque les canaux sont supposés être i.i.d. de bloc en bloc. Lorsque $L \to \infty$, la contrainte devient de moins en moins restrictive et est donc toujours satisfaite. Toute distribution sur $\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2$ est donc atteignable.

Finalement, il nous paraît particulièrement important de souligner que la contrainte d'information obtenue en faisant de la coordination via les actions est différente de ce qui serait obtenu en faisant de la coordination via la communication classique [79] avec un canal dédié. Si l'Agent 1 pouvait communiquer avec l'Agent 2 via un canal de capacité C, alors toutes distributions conjointes de probabilité $\widetilde{Q} \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ qui vérifie la contrainte d'information

$$I_{\widetilde{Q}}(X_0; X_2) \leq C \tag{2.28}$$

serait atteignable. En revanche, la contrainte $I_Q(X_0; X_2) \leq I_Q(X_1; Y | X_0, X_2)$ reflète les deux caractéristiques suivantes de la communication via les actions :

- la distribution de ce que nous appelons les entrées X_1 du canal implicite (pour reprendre une terminologie traditionnelle) entre l'Agent 1 et l'Agent 2 ne peut pas être optimisée séparément des actions de l'Agent 2 et de l'état. Au contraire, X_1 et (X_0, X_2) doivent être corrélés d'une manière appropriée, en fonction de la fonction d'utilité d'étape considérée ;

- Pour une approche classique, les réalisations de Y, les sorties du canal dans la terminologie traditionnelle, ne dépendent que des réalisations des entrées du canal X_1 alors qu'ici ils dépendent également de (X_0, X_2). L'état X_0 et les actions X_2 jouent le rôle de l'état du canal implicite.

Sous des conditions spécifiques, la coordination via les actions peut se réduire à de la coordination via un canal de communication dédié. Par exemple, si la fonction d'utilité se factorise comme suit :

$$w(x_0, x_1, x_2) \triangleq w_1(x_1) w_2(x_0, x_2), \tag{2.29}$$

et si la structure d'observation satisfait la chaîne de Markov $X_0, X_2 - X_1 - Y$, alors toute distribution $\widetilde{Q}(x_0, x_1, x_2) \triangleq \widetilde{Q}(x_0, x_2)\widehat{Q}(x_1)$ qui satisfait la contrainte d'information

$$I_{\widetilde{Q}}(X_0; X_2) \leq I_{\widehat{Q}}(X_1; Y) \tag{2.30}$$

sera atteignable. En particulier, il est possible d'optimiser $\widehat{Q}$ de manière indépendante. De plus, si $w_1(x_1)$ est une fonction constante, la contrainte d'information se simplifie :

$$I_{\widetilde{Q}}(X_0; X_2) \leq \max_{\widehat{Q}} I_{\widehat{Q}}(X_1; Y), \tag{2.31}$$

et le canal de communication implicite devient effectivement un canal dédié.

Atteignabilité : Cas II

Nous considérons maintenant le cas II, c'est-à-dire le cas correspondant aux stratégies (2.7) pour les agents.

Théorème 2.8. *Considérons la structure d'observation du Cas II. Soit* $\overline{Q} \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ *avec* $\sum_{(x_1,x_2)} \overline{Q}(x_0, x_1, x_2) = \rho_0(x_0)$, $\forall x_0 \in \mathcal{X}_0$. *Si la distribution de probabilité* $Q \in \Delta(\mathcal{U} \times \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})$, *structurée comme suit*

$$Q(u, x_0, x_1, x_2, y) = \mathrm{P}(u|x_0, x_1, x_2)\Gamma(y|x_0, x_1, x_2) \cdot \overline{Q}(x_0, x_1, x_2), \\ \forall\, (u, x_0, x_1, x_2, y) \in \mathcal{U} \times \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y} \tag{2.32}$$

vérifie la contrainte :

$$I_Q(X_0; X_2) \leq I_Q(U; Y, X_2) - I_Q(U; X_0, X_2), \tag{2.33}$$

alors $\overline{Q}$ *est atteignable au sens de la Définition 2.2, pour des stratégies de la forme (2.7).*

La démonstration de ce théorème reprend les mêmes principes que celle du Théorème 2.7, mais est cependant plus complexe, notamment à cause de la présence d'une *variable auxiliaire* U. En plus de la méthode des types, nous utilisons la technique de « random binning » [81]. Cette technique divise le livre de code (canal dans notre cas) en paquets (les bins), en affectant aléatoirement chaque mot de code canal à un paquet. C'est cette configuration qui fera apparaître la *variable auxiliaire*. Cette technique permet de réduire en partie l'influence du manque d'observations de l'Agent 2 qui, nous le rappelons, n'observe rien sur la séquences d'états.

Quelques remarques importantes sont à faire ici :

- Même si nous exploitons un schéma classique de codage source canal séparé, une différence clé doit être notée par rapport au codage classique de Gel'fand et Pinsker [42] et ses extensions (par exemple [66]). En effet, même si utiliser les actions passées du décodeur (l'Agent 2) n'aide pas à augmenter la capacité du canal, cela est utile lorsque nous voulons faire de la coordination. Le codage classique de Gel'fand et Pinsker donnerait la contrainte d'information

  $$I_Q(X_0; X_2) \leq I_Q(U; Y) - I_Q(U; X_0, X_2) \tag{2.34}$$

  qui est plus restrictive que (2.33).

- Dans le cas particulier où l'Agent 2 **observe parfaitement** les actions de l'Agent 1, c'est-à-dire lorsque $Y = X_1$, nous pouvons remarquer que (2.19), (2.24), (2.33), et (2.34) (pour cette contrainte, prendre $U = X_1$) coïncident. Elles sont toutes égales à

$$I_Q(X_0; X_2) \leq H_Q(X_1|X_0, X_2) \tag{2.35}$$

  Cela implique que connaître X_0 de manière strictement causale pour l'Agent 2 (comme dans le cas II) n'apporte aucun gain en terme de performance.

- Nous n'avons pas prouvé que la contrainte d'information du Théorème 2.8 est une condition nécessaire d'atteignabilité pour le cas II. Ici aussi, il pourrait être tentant d'utiliser une interprétation en terme d'information adjacente pour démontrer la réciproque. En effet, (2.33) peut être lié au cas de Cover et Chiang [25] où l'état du canal est $S = (S_1, S_2)$, et S_1 (respectivement S_2) est connue à l'encodeur (respectivement au décodeur). Mais comme expliqué précédemment, réutiliser les démonstrations existantes de réciproque ne fonctionne pas et trouver la variable auxiliaire appropriée ne semble pas trivial.

Dans la section suivante, nous réécrivons le problème comme un problème de minimisation sous contraintes, cette formulation nous étant très utile pour l'application considérée en Section 2.5.

2.4 Problème d'optimisation sous contraintes

L'objectif de cette section est d'étudier le problème d'optimisation associé à la détermination de la meilleure distribution conjointe de probabilité Q pour les cas I et II. Le problème sera formulé sous deux formes différentes, une impliquant Q (qui sera vue comme une fonction, comme définie dans les sections précédentes), et une deuxième sous forme vectorielle, impliquant le vecteur $q^{n^\ell} = (q_1, q_2, ..., q_{n^\ell})$, $\ell \in \{\mathrm{I}, \mathrm{II}\}$, avec $n^{\mathrm{I}} = |\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}|$ et $n^{\mathrm{II}} = |\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y} \times \mathcal{U}|$. Cette deuxième formulation impliquera des notations assez complexes, mais est très adaptée à l'implémentation numérique et donc à la détermination effective de l'utilité espérée maximum. Ce qui est d'un intérêt central, puisque cela donne la valeur, et non uniquement la caractérisation analytique, de l'utilité espérée du problème sous contraintes d'information que nous considérons. Pour aller plus loin que dans la Section 2.4.1 en terme de connaissance sur la structure de la solution, nous étudions en Section 2.4.2 le cas particulier d'**observation parfaite**. Rappelons que les contraintes d'information (2.24) et (2.33) coïncident dans ce cas. Les problèmes d'importances que nous étudions sont la relation entre la fonction d'utilité d'étape et la (les) meilleure(s) distribution(s) conjointe(s) de probabilité, ainsi que l'unicité de la solution.

2.4.1 Cas général

En exploitant les résultats de la Section 2.3, nous voyons qu'une distribution conjointe optimale est une solution du problème d'optimisation sous contraintes suivant, pour $\ell \in \{\mathrm{I}, \mathrm{II}\}$:

$$
\begin{aligned}
\text{Minimiser} \quad & -\mathbb{E}_Q[w(X)] = -\sum_{x,y,u} Q(x,y,u)w(x) & & \\
\text{sous contraintes :} \quad & -1 + \sum_{(x,y,u)} Q(x,y,u) & \overset{(c)}{=} & \ 0 \\
\forall (x,y,u), \quad & \frac{Q(x,y,u)}{\sum_{(y,u)} Q(x,y,u)} - \Gamma(y|x) & \overset{(d)}{=} & \ 0 \\
\forall x_0, \quad & -\rho_0(x_0) + \sum_{x_1,x_2,y,u} Q(x,y,u) & \overset{(e)}{=} & \ 0 \\
\forall (x,y,u), \quad & -Q(x,y,u) & \overset{(f)}{\leq} & \ 0 \\
& \Phi^{\ell}(Q) & \overset{(g)}{\leq} & \ 0
\end{aligned}
\tag{2.36}
$$

où $X = (X_0, X_1, X_2)$, $x = (x_0, x_1, x_2)$, et dans le Cas I, $\Phi^{\mathrm{I}}(\cdot)$ est définie en équation (2.20) alors que dans le cas II :

$$
\Phi^{\mathrm{II}}(Q) \triangleq I_Q(X_0; X_2) - I_Q(U; Y, X_2) + I_Q(U; X_0, X_2) \tag{2.37}
$$

Dans le cas I, l'ensemble $\mathcal{U}$ se réduit à un singleton.

Une des premières questions à se poser est de savoir si ce problème est convexe par rapport à Q [17]. Nous avons :

- La fonction objectif est linéaire en Q ;
- Le simplexe unité, défini par (c) et (f) est un ensemble convexe ;
- Les contraintes d'égalité (d) et (e) définissent un ensemble convexe ;

Il reste donc à prouver que l'ensemble défini par l'inégalité de contrainte d'information (g) est convexe pour que le problème (2.36) soit convexe. Nous savons déjà via le Lemme 2.6 que pour le cas I, la fonction Φ^{I} est convexe par rapport à Q, ce qui prouve que l'ensemble défini par l'équation (g) est convexe. Dans le cas II, nous n'avons pas prouvé que Φ^{II} est une fonction convexe mais en utilisant un argument de partage du temps [22], il est possible de convexifier l'ensemble correspondant. Nous supposerons que c'est le cas dans la suite de ce chapitre.
Comme le problème est convexe, une deuxième question importante est de savoir s'il existe un point strictement faisable, ce qui impliquera que la **condition de Slater** est vérifiée et que les **conditions de Karush-Kuhn-Tucker (KKT)** sont des conditions nécessaires d'optimalité. Le problème étant convexe, elles sont aussi suffisantes.
Pour le cas II, un point strictement faisable existe effectivement, comme le prouve la proposition suivante. La preuve pour le cas I est un cas particulier.

Proposition 2.9. *La* ***condition de Slater*** *est vérifiée dans les cas I et II pour des probabilités de transition du canal* irréductibles*, c'est-à-dire telles que*

$$
\forall\, (x_0, x_1, x_2, y) \in \mathcal{X} \times \mathcal{Y} : \qquad \Gamma(y|x_0, x_1, x_2) > 0 \tag{2.38}
$$

22. « time-sharing » en anglais.

En effet, en prenant des variables aléatoires X_0, X_1, et X_2 indépendantes, et en supposant que $U = X_1$, il est facile de prouver que cette condition est vérifiée.

Maintenant, nous voulons réécrire le problème d'optimisation ci-dessus en utilisant le vecteur $q^{n^\ell} = (q_1, q_2, ..., q_{n^\ell})$ plutôt que la fonction Q. Cela sera à la fois utile et nécessaire pour utiliser les solveurs numériques standards, ce qui sera fait en Section 2.5, et pour appliquer les conditions KKT, ce qui sera fait en Section 2.4.2. Pour réécrire le problème, il nous faut introduire quelques notations. La cardinalité de l'ensemble $\mathcal{X}_k$, $k \in \{0, 1, 2\}$, sera dénotée $n_i \triangleq |\mathcal{X}_k| < \infty$. Nous supposerons donc, par soucis de simplicité, que $\mathcal{X}_k$ est défini comme un ensemble d'indices $\mathcal{X}_k = \{1, ..., n_k\}$. Nous utilisons des notations similaires pour $\mathcal{U} = \{1, \dots, n_u\}$ et $\mathcal{Y} = \{1, \dots, n_y\}$, $n_y \triangleq |\mathcal{Y}|$ et $n_u \triangleq |\mathcal{U}|$. Avec cette convention, nous définissons la fonction bijective $\psi^\ell : \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y} \times \mathcal{U} \to \{1, \cdots, n^\ell\}$ comme

$$\psi^\ell(i', j', k', l', m') \triangleq m' + n_u(l' - 1) + n_u n_y(k' - 1) + n_u n_y n_2(j' - 1) + n_u n_y n_2 n_1(i' - 1), \tag{2.39}$$

qui met en relation une réalisation $(i', j', k', l', m') \in \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y} \times \mathcal{U}$ avec un unique indice $\psi^\ell(i', j', k', l', m') \in \{1, \cdots, n^\ell\}$.
Nous notons également $n^{\mathrm{I}} \triangleq n_0 n_1 n_2 n_y$ and $n^{\mathrm{II}} \triangleq n_0 n_1 n_2 n_y n_u$. Ce nous permet d'introduire le vecteur de probabilité $q^{n^\ell} = (q_1, q_2, \cdots, q_{n^\ell})$ pour $\ell \in \{\mathrm{I}, \mathrm{II}\}$, pour lequel chaque composant q_i, $i \in \{1, \cdots, n^\ell\}$, est égal à $Q((\psi^\ell)^{-1}(i))$, et le vecteur de valeurs de paiement $w^{n^\ell} = (w_1, w_2, \cdots, w_{n^\ell}) \in \mathbb{R}^{n^\ell}$, pour lequel chaque composant w_i est le paiement de $(\psi^\ell)^{-1}(i)$. La relation entre la fonction Q (respectivement w) et le vecteur q^{n^ℓ} (respectivement w^{n^ℓ}) est résumé dans le Tableau 2.3.

En utilisant cette indexation, le problème d'optimisation peut être réécrit sous la forme standard suivante :

$$\begin{aligned}
\text{Minimiser} \quad & -\mathbb{E}_Q[w(X)] = -\sum_{i=1}^{n^\ell} q_i w_i & & \\
\text{sous contraintes :} \quad & -1 + \sum_{i=1}^{n^\ell} q_i & \overset{(h)}{=} & 0 \\
\forall i \in \{1, \dots, n^\ell\} \quad & \frac{q_i}{\Theta_i} - \Gamma_i & \overset{(i)}{=} & 0 \\
\forall i \in \{1, \dots, n_0\} \quad & -\rho_{0i} + \sum_{j=1+(i-1)n_1 n_2 n_y n_u}^{i n_1 n_2 n_y n_u} q_j & \overset{(j)}{=} & 0 \\
\forall i \in \{1, ..., n^\ell\} \quad & -q_i & \overset{(k)}{\leq} & 0 \\
& \phi^\ell(q^{n^\ell}) & \overset{(\ell)}{\leq} & 0
\end{aligned} \tag{2.40}$$

où

$$\Theta_i \triangleq \sum_{\substack{j \in \{1, \dots, n_y n_u\} \\ k \in \{1, \dots, n_0 n_1 n_2\}}} q_{(k-1)n_y n_u + j} \cdot \mathbb{1}_{\{(k-1)n_y n_u \leq i \leq k n_y n_u - 1\}}, \tag{2.41}$$

$\forall\ i \in \{1, ..., n_0\}$, ρ_{0i} correspond à $\mathrm{P}(X_0 = i) = \rho_0(i)$, $\forall\ i \in \{1, ..., n^\ell\}$, et Γ_i correspond à la valeur de $\Gamma(y|x_0, x_1, x_2)$, par rapport au Tableau 2.3. Pour la fonction associée à la contrainte (e), elle s'écrit dans le cas II (le cas I suit par spécialisation à $|\mathcal{U}| = 1$) comme suit :

Indice de q_i	X_0	X_1	X_2	Y	U
1	1	1	1	1	1
2	1	1	1	1	2
⋮	⋮	⋮	⋮	⋮	
n_u	1	1	1	1	n_u
n_u+1	1	1	1	2	1
⋮	⋮	⋮	⋮	⋮	⋮
$2n_u$	1	1	1	2	n_u
⋮	⋮	⋮	⋮	⋮	⋮
$n_1n_2n_yn_u-n_u+1$	1	n_1	n_2	n_y	1
⋮	⋮	⋮	⋮	⋮	⋮
$n_1n_2n_yn_u$	1	n_1	n_2	n_y	n_u
⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮
$(n_0-1)n_1n_2n_yn_u+1$	n_0	1	1	1	1
⋮	⋮	⋮	⋮	⋮	⋮
$n_0n_1n_2n_yn_u$	n_0	n_1	n_2	n_y	n_u

Tableau 2.3 – Indexation choisie pour le vecteur d'utilité w et pour le vecteur de distribution de probabilité q. Les lignes en gras délimitent un bloc de taille $n_1n_2n_yn_u$ et chaque bloc correspond à une valeur donnée de l'état aléatoire X_0. Les 5−uplets sont ordonnés par rapport à l'ordre lexicographique.

$$\phi^{\mathrm{II}}(q^{n^{\mathrm{II}}}) = I_{q^{n^{\mathrm{II}}}}(X_0;X_2) - I_{q^{n^{\mathrm{II}}}}(U;Y,X_2) + I_{q^{n^{\mathrm{II}}}}(U;X_0,X_2) \tag{2.42}$$

$$= H_{q^{n^{\mathrm{II}}}}(X_0) - H_{q^{n^{\mathrm{II}}}}(U,X_0|X_2) + H_{q^{n^{\mathrm{II}}}}(U|Y,X_2) \tag{2.43}$$

$$= H_{q^{n^{\mathrm{II}}}}(X_0) + H_{q^{n^{\mathrm{II}}}}(X_2) - H_{q^{n^{\mathrm{II}}}}(X_0,X_2,U) + H_{q^{n^{\mathrm{II}}}}(X_2,Y,U) - H_{q^{n^{\mathrm{II}}}}(X_2,Y) \tag{2.44}$$

où les définitions de chaque terme peuvent être trouvées dans l'article de la section E.1.1 en Annexe E.

Cette formulation est directement utilisée dans la sous-section suivante ainsi que dans la Section 2.5.

2.4.2 Le cas de l'observation parfaite $Y = X_1$

Comme nous l'avons déjà dit, en observation parfaite, (2.24) et (2.33) coïncident :

$$\phi(q^n) \triangleq \phi^{\mathrm{I}}(q^{n^{\mathrm{I}}}) = \phi^{\mathrm{II}}(q^{n^{\mathrm{II}}})$$
$$= H_{q^n}(X_0) + H_{q^n}(X_2|X_0) - H_{q^n}(X_1|X_0,X_2) \tag{2.45}$$

avec $q^n = (q_1,...,q_n)$, $n = n_0n_1n_2$. Pour mieux comprendre la relation entre le vecteur de valeurs d'utilité w^n et la distribution conjointe de probabilité optimale q^n, nous exprimons les conditions

de KKT. Le **Lagrangien** du problème s'écrit :

$$\mathcal{L}(q^n, \lambda^n, \mu_0, \mu^{n_0}, \lambda_{\text{IC}}) = -\sum_{i=1}^{n} w_i q_i + \lambda_i q_i + \mu_0 \left[-1 + \sum_{i=1}^{n} q_i\right] + \sum_{i=1}^{n_0} \mu_i \left[-\rho_{0i} + \sum_{j=1+(i-1)n_1n_2}^{in_1n_2} q_j\right] + \lambda_{\text{CI}}\phi(q^n) \tag{2.46}$$

où $\lambda^n \triangleq (\lambda_1, ..., \lambda_n)$, $\mu^{n_0} \triangleq (\mu_1, ..., \mu_{n_0})$, μ_0, et λ_{CI} sont les multiplicateurs de Lagrange et CI signifiant Contrainte d'Information.

Une condition nécessaire et suffisante pour qu'une distribution de probabilité q^n soit un point optimal est que cette distribution soit solution du système suivant :

$$\frac{\partial \mathcal{L}}{\partial q_i} = -w_i - \lambda_i + \mu_0 + \sum_{j=1}^{n_0} \mu_j \mathbb{1}_{\{1+n_1n_2(j-1) \le i \le jn_1n_2\}} + \lambda_{\text{IC}} \frac{\partial \phi}{\partial q_i}(q^n) = 0 \; \forall \, i \in \{1, \ldots, n\} \tag{2.47}$$

$$(h), (i), (j) \tag{2.48}$$

$$\lambda_i \ge 0 \qquad \forall \, i \in \{1, \ldots, n\} \tag{2.49}$$

$$\lambda_{\text{CI}} \ge 0 \tag{2.50}$$

$$\lambda_i q_i = 0 \qquad \forall \, i \in \{1, \ldots, n\} \tag{2.51}$$

$$\lambda_{\text{CI}}\phi(q^n) = 0 \tag{2.52}$$

où

$$\frac{\partial \phi}{\partial q_i}(q^n) = \left[-\sum_{k=1}^{n_0} \left(\mathbb{1}_{\{1+(k-1)n_1n_2 \le i \le kn_1n_2\}} \log \sum_{j=1+(k-1)n_1n_2}^{kn_1n_2} q_j \right) - \sum_{k=1}^{n_2} \mathbb{1}_{\{i \in \{k, k+n_2, \ldots, k+(n_0n_1-1)n_2\}\}} \log \sum_{j=0}^{n_0n_1-1} q_{k+jn_2} - 1 + \log q_i \right] \; \forall i \in \{1, \ldots, n\}. \tag{2.53}$$

D'après les conditions KKT, quelques commentaires peuvent être fait, spécialement quand le vecteur de valeurs d'utilité w^n peut être strictement ordonné, ce que nous supposons dans la suite :

- Premièrement, si l'utilité espérée était maximisée sour les contraintes (h) et (k) uniquement, la meilleure distribution conjointe serait d'associer une probabilité 1 au plus grand élément du vecteur w^n. Ainsi, seulement le triplet (x_0, x_1, x_2) associé à l'utilité maximale serait utilisée. La meilleure q^n correspondrait donc à un sommet du simplexe unité $\Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$.

- Deuxièmement, et comme la distribution de probabilité de l'état aléatoire est donnée (contrainte (j)), au moins n_0 composantes de q^n doivent être strictement positives. Nous pouvons vérifier que sous les contraintes $(h), (j)$, et (k), la solution optimale est telle que, pour tout x_0, la paire optimale (x_1, x_2) est choisie. q^n aura donc exactement n_0 composantes strictement positives. Mais ceci ne prend pas en compte la présence de la contrainte

d'information qui limite les capacités de communication entre les agents. La présence de cette contrainte (ℓ) implique que, en général, plus que n_0 composantes sont positives pour la (les) solution(s) optimale(s).

- Troisièmement, sous l'hypothèse que le vecteur d'utilité peut être strictement ordonné, la solution est unique :

  Proposition 2.10. *S'il existe une permutation telle que le vecteur d'utilité w^n peut être strictement ordonné, alors le problème d'optimisation* (2.40) *a une unique solution.*

  La preuve de cette proposition utilise notamment le lemme de convexité 2.6. Le point technique est de prouver que le multiplicateur de Lagrange λ_{CI} associé à la contrainte d'information est strictement positive, rendant la contrainte d'information active.

Nous passons maintenant à l'application aux problèmes de contrôle de puissance.

2.5 Application : le *contrôle de puissance codé*

2.5.1 Description du cadre proposé

A partir de maintenant, nous considérons un exemple particulier du modèle développé dans les sections précédentes : le **problème de contrôle de puissance dans les réseaux avec interférence**. Dans ce contexte, les agents sont des *émetteurs* et l'état du système est donné par l'*état global du canal* c'est-à-dire tous les gains des canaux associés aux différents liens entre émetteurs et récepteurs. L'ensemble $\mathcal{X}_0$ représente donc l'alphabet de l'état global du canal.
L'idée du **contrôle de puissance codé** consiste à intégrer l'information à propos de l'état du canal dans les niveaux de puissance de transmission eux-mêmes. Dans cette nouvelle approche, en supposant que les niveaux de puissance d'un émetteur donné peuvent être observés par les autres émetteurs via un certain canal, la séquence de niveaux de puissance du premier peut être utilisée pour se coordonner avec les autres. Le canal d'observation sous-jacent peut typiquement correspondre à un mécanisme de détection (comme pour les paramètres de radio cognitive) pour un mécanisme de retour d'information [23] (comme supposé classiquement dans les réseaux à interférences). Une des principales caractéristiques du contrôle de puissance codé est que les interférences peuvent être gérées directement dans le domaine de fréquence radio et n'exige pas de détection de « baseband » ou de décodage, ce qui est très utile par exemple dans les réseaux hétérogènes. Créer un canal de signalisation dédié pour coordonner les émetteurs n'est donc pas nécessaire.

Le but principal de cette section est d'évaluer les performances limites du contrôle de puissance codé et les gains potentiels qu'il peut amener par rapport aux approches classiques telles que l'implémentation d'une stratégie de contrôle de puissance type Nash pour un jeu non-coopératif donné. Pour cela, un canal à interférence avec deux émetteurs et deux récepteurs est considéré. Dans la section 2.5.5, le cas particulier du canal à accès multiple est étudié. En particulier, il nous permet de pourvoir un code de contrôle de puissance non trivial que l'on peut interpréter plus facilement.

23. « feedback »

En dénotant g_{ij} le gain du canal entre l'Emetteur i et le Recepteur j, $i,j \in \{1,2\}$, chaque réalisation de l'état du canal est donnée par :

$$x_0 = (g_{11}, g_{12}, g_{21}, g_{22}), \tag{2.54}$$

où $g_{ij} \in \mathcal{G}$, $|\mathcal{G}| < \infty$. Nous supposons que les gains du canal g_{ij} sont indépendants. L'alphabet $\mathcal{X}_i, |\mathcal{X}_i| < \infty$, $i \in \{1,2\}$, représente l'ensemble des niveaux de puissance possible pour l'Emetteur i. Notons que supposer que ces ensembles $\mathcal{X}_i, i \in \{0,1,2\}$ sont discrets a un intérêt pratique. Par exemple, il existe des standards de communication sans fil pour lesquels la puissance peut être augmenter ou diminuer uniquement par une certaine valeur de puissance, et l'information sur l'état du canal est quantifiée (par exemple dans certains systèmes de réseaux locaux cellulaires ou sans fil). De plus, utiliser des niveaux de puissance discrets peut ne pas induire de perte d'optimalité [43]. Nous reviendrons sur ce point plus en détail en Section 2.5.2.

Pour les fonctions d'utilité d'étape, nous considérerons pas moins de trois fonctions différentes :

$$w^{\text{taux}} : \left| \begin{array}{ccc} \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 & \to & \mathbb{R}_+ \\ (x_0, x_1, x_2) & \mapsto & \displaystyle\sum_{i=1}^{2} \log_2 \left(1 + \frac{g_{ii}x_i}{\underbrace{\sigma^2 + g_{-ii}x_{-i}}_{\text{RSIB}_i}} \right) \end{array} \right. , \tag{2.55}$$

$$w^{\text{RSIB}} : \left| \begin{array}{ccc} \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 & \to & \mathbb{R}_+ \\ (x_0, x_1, x_2) & \mapsto & \displaystyle\sum_{i=1}^{2} \frac{g_{ii}x_i}{\sigma^2 + g_{-ii}x_{-i}} \end{array} \right. , \tag{2.56}$$

$$w^{\text{energie}} : \left| \begin{array}{ccc} \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 & \to & \mathbb{R}_+ \\ (x_0, x_1, x_2) & \mapsto & \displaystyle\sum_{i=1}^{2} \frac{F\left(1 + \frac{g_{ii}x_i}{\sigma^2 + g_{-ii}x_{-i}}\right)}{x_i} \end{array} \right. \tag{2.57}$$

où :

- $\mathcal{X}_0 = \mathcal{G}^4$;
- les fonctions d'utilité d'étape w^{taux}, w^{RSIB}, et w^{energie} représente respectivement la somme des taux, la somme des RSIB, et la somme des efficacités énergétiques ;
- la notation $-i$ signifie l'émetteur autre que i ;
- σ^2 correspond au niveau de bruit à la réception ;
- $F : \mathbb{R}_+ \to [0,1]$ est une fonction sigmoïdale et croissante qui représente généralement le taux de succès par bloc (voir par exemple [70][10] pour plus de détail) ;
- la fonction F est choisie telle que w^{energie} est continue est a une limite lorsque $x_i \to 0$;
- enfin, les structures d'observation considérées seront les trois suivantes :
  - Observation parfaite : $Y = X_1$;

– Observation CBS : la structure d'observation correspond à un *canal binaire symétrique* (CBS), c'est-à-dire $Y = X_1 \oplus Z_1$ où X_1, Y, et Z_1 sont des variables aléatoires de Bernoulli ;
– Retour d'information bruité du RSIB : la Figure 2.4 représente le modèle de canal avec retour d'information que nous considérons dans ce chapitre. L'Emetteur 2 peut uniquement voir les niveaux de puissance de l'Emetteur 1 à travers une version bruitée de son RSIB.

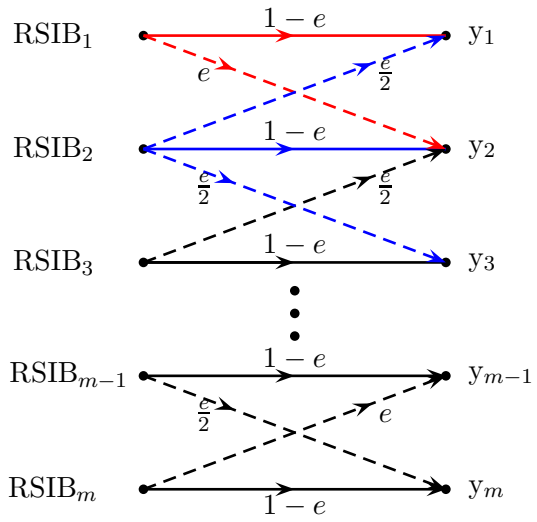

FIGURE 2.4 – **Canal avec retour d'information RSIB** : le signal observé par l'Emetteur 2 appartient à un alphabet à m symboles, $|\mathcal{Y}| = m$. Les symboles correspondent aux différentes valeurs possibles pour le RSIB au récepteur qui est associé à l'Emetteur 2. L'Emetteur 2 observe la valeur réelle de son RSIB avec probabilité $1 - e$, il y a donc une probabilité $e \geq 0$ qu'un RSIB voisin soit observé. Dans les simulations, nous prenons $e = 0.1$.

Quand les gains relatifs sont pourvus, nous considérons une des trois stratégies de référence en contrôle de puissance :

- La stratégie de contrôle de puissance Nash (CPN) : elle suppose que chaque émetteur veut maximiser sa fonction d'utilité d'étape individuelle $u_i(x_0, x_1, x_2)$. Pour la somme des taux, la somme des RSIB et la somme des efficacités énergétiques, ces fonctions d'utilité d'étape individuelles sont données respectivement par

$$u_i(x_0, x_1, x_2) = \log_2(1 + \mathrm{RSIB}_i), \tag{2.58}$$

$$u_i(x_0, x_1, x_2) = \mathrm{RSIB}_i, \tag{2.59}$$

et

$$u_i(x_0, x_1, x_2) = \frac{F(\mathrm{RSIB}_i)}{x_i}. \tag{2.60}$$

Dans les cas somme des taux et somme des RSIB, l'unique équilibre de Nash est $(x_1^{\mathrm{NE}}, x_2^{\mathrm{NE}}) = (P_{\max}, P_{\max})$, quelque soit la valeur de x_0. Dans le cas de l'efficacité énergétique, l'unique

équilibre de Nash (non trivial) peut être déterminé numériquement et requiert généralement une certaine connaissance de x_0, en fonction de la façon dont il est implémenté.

- La stratégie de contrôle de puissance semi-coordonné (CPS) : $x_2 = P_{\max}$ et

$$x_1^\dagger \in \arg\max_{x_1} w^{\mathrm{r}}(x_0, x_1, P_{\max}) \qquad \mathrm{r} \in \{\text{taux}, \text{RSIB}, \text{energie}\} \tag{2.61}$$

  CPS correspond au scénario où l'Emetteur 1 est informé que l'Emetteur 2 transmet à puissance maximale. CPS n'a besoin que de la connaissance de la réalisation de l'état courant à l'Emetteur 1.

- Le cas de communication sans coût (déjà évoqué en Section 2.1.1, équation (2.3)) où le maximum de w^{r} peut être atteint à chaque étape. Cette stratégie donne une borne supérieure sur la performance. Si l'état est i.i.d. de bloc en bloc (c'est-à-dire que l'état est constant sur toutes les étapes d'un même bloc), et le temps de cohésion augmente, la borne tend à être atteinte.

Le rapport signal sur bruit (RSB)[24] est défini par :

$$\mathrm{RSB(dB)} = 10 \log_{10} \frac{P_{\max}}{\sigma^2}. \tag{2.62}$$

2.5.2 Influence de la fonction d'utilité

Le but de cette section est de fournir les gains de performance relatifs (en %) qui peuvent être obtenus avec le contrôle de puissance codé par rapport au schéma CPS, qui est déjà une stratégie de coordination. Pour cela, nous supposons que les gains de canaux suivent une loi de Bernoulli :

$$g_{ij} \sim \mathcal{B}(p_{ij}), \tag{2.63}$$

$$g_{ij} \in \{g_{\min}, g_{\max}\} \tag{2.64}$$

avec $\mathrm{P}(g_{ij} = g_{\min}) = p_{ij}$. Ce qui implique que $|\mathcal{X}_0| = 16$.
De plus, dans cette section, nous prenons :

- $g_{\min} = 0.1$;
- $g_{\max} = 2$;
- $(p_{11}, p_{12}, p_{21}, p_{22}) = (0.5, 0.1, 0.1, 0.5)$;
- les ensembles de puissance de transmission $\mathcal{X}_1, \mathcal{X}_2$ sont supposés être des alphabets à quatre éléments $\{P_1, P_2, P_3, P_4\}$ avec $P_1 = 0, P_2 = \frac{P_{\max}}{3}, P_3 = \frac{2P_{\max}}{3}, P_4 = P_{\max}$;
- $P_{\max}$ est donné par le RSB effectif ;
- $\sigma^2 = 1$;

24. « Signal to Noise Ratio » (SNR)

- la fonction F est choisie comme une fonction d'efficacité typique, utilisée notamment dans [10] :

$$F(x) = \exp\left(-\frac{2^{0.9}-1}{x}\right). \tag{2.65}$$

Enfin, l'**observation parfaite** est supposée, c'est-à-dire $Y = X_1$.
Pour tout r $\in \{\text{rate}, \text{SINR}, \text{energie}\}$, nous définissons les gains relatifs par rapport à la stratégie CPS par :

$$\left(\frac{\mathbb{E}_{\overline{Q}^\star}(w^{\mathrm{r}})}{\mathbb{E}_{\rho_0}(\max\limits_{x_1} w(x_0, x_1, P_{\max}))} - 1\right) \times 100 \tag{2.66}$$

où $\overline{Q}^\star$ est la distribution conjointe de probabilité qui est trouvée en résolvant le problème d'optimisation (2.40) en le particularisant au cas d'observation parfaite. Cette optimisation est réalisée via le programme Matlab fmincon [25]. La Figure 2.5 représente les gains relatifs de performance par rapport à la stratégie CPS pour la somme des efficacités énergétiques alors que la Figure 2.6 correspond à la somme des RSIB et à la somme des taux. Les gains les plus importants amenés par le contrôle de puissance codé sont obtenus pour la somme des efficacités énergétiques. Cette fonction est particulièrement sensible au manque de coordination. En effet, quand la puissance de transmission augmente, $\frac{F(\text{SINR}_i)}{x_i} \to \frac{1}{x_i}$, ce qui veut dire que l'efficacité énergétique diminue très rapidement. Ce qui explique les gains relatifs très importants que nous obtenons ici. Des gains relatifs jusqu'à 43% peuvent être obtenus pour la somme des RSIB (voir la courbe du haut sur la Figure 2.6). Enfin, les gains pour le cas somme des taux (voir la courbe du bas sur la Figure 2.6) sont moins importants mais quand même assez significatif : jusqu'à 25%.
Notons que tous ces gains pourraient être bien plus grand si le cas de référence était la stratégie CPN comme cela est souvent considéré dans la littérature sur le contrôle de puissance distribué.
Pour conclure cette section, le Tableau 2.7 fournit les distributions marginales

$$\overline{Q}^\star_{X_1}(x_1) = \sum_{x_0,x_2} \overline{Q}^\star(x_0, x_1, x_2),\ \overline{Q}^\star_{X_2}(x_2) = \sum_{x_0,x_1} \overline{Q}^\star(x_0, x_1, x_2), \tag{2.67}$$

et la distribution conjointe

$$\overline{Q}^\star_{X_1X_2}(x_1, x_2) = \sum_{x_0} \overline{Q}^\star(x_0, x_1, x_2) \tag{2.68}$$

dans le cas de la somme des taux pour un point optimal (pour le problème d'optimisation sous contraintes), et pour RSB = 10dB. Le Tableau 2.8 fournit quand à lui la même information pour le cas de communication sans coût (c'est-à-dire sans la contrainte d'information). Ce tableau montre que, sans contrainte d'information, la somme des taux est maximisée lorsque les émetteurs corrèlent leurs niveaux de puissance de telle façon que seulement trois paires de niveaux de puissance de transmission sont utilisées (sur 16). Ce résultat a été prouvé dans [43]. Dans cet article, il est démontré que pour les canaux à interférences avec deux paires d'émetteurs/récepteurs, il n'y a aucune perte d'optimalité en terme de w^{taux} lorsque nous passons d'un intervalle de puissance disponible continu $[0, P_{\max}]$ à un ensemble binaire $\{0, P_{\max}\}$. Un dernier point intéressant, comme l'indique le Tableau 2.7, est que les trois meilleures configurations d'interférence pour le cas de communication sans coût sont exploitées pour le cas sous contraintes $44.3 + 42.9 + 2.1 = 89.3\%$ du temps malgré la présence de ces contraintes de communication entre les émetteurs.

25. Une descprition détaillé de ce programme peut être trouvé à l'adresse suivante : http://www.mathworks.fr/fr/help/optim/ug/fmincon.html.

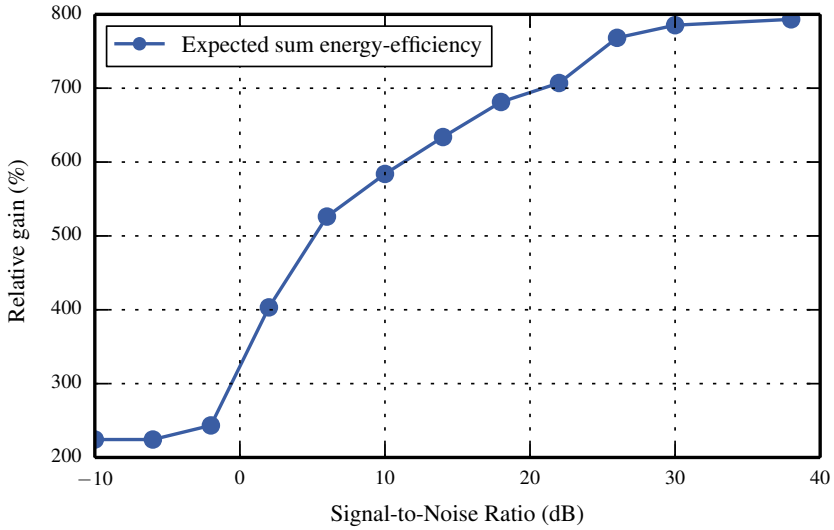

FIGURE 2.5 – **Gains relatifs, cas somme des efficacités énergétiques** : Gains relatifs en % (par rapport à la stratégie CPS) en terme d'utilité espérée pour la somme des efficacités énergétiques par rapport au Rapport Signal sur Bruit (RSB).

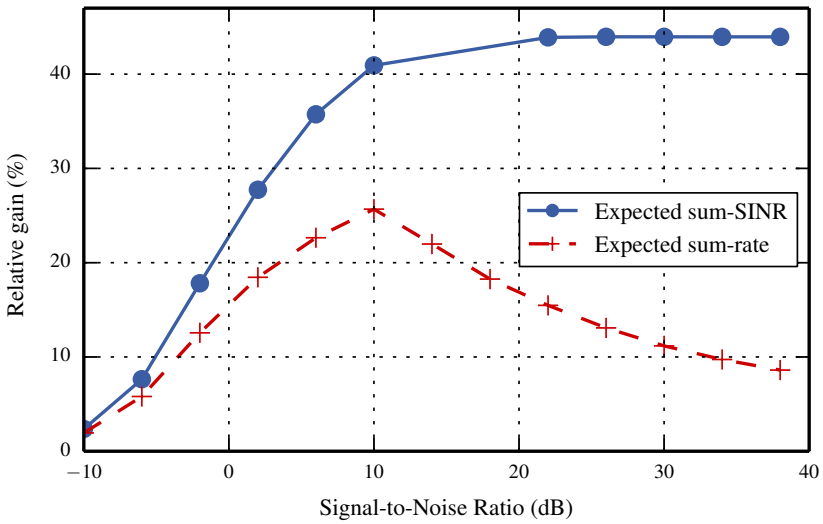

FIGURE 2.6 – **Gains relatifs, cas somme des taux et somme des RSIB** : Gains relatifs en % (par rapport à la stratégie CPS) en terme d'utilité espérée pour la somme des taux (courbe du bas) et la somme des RSIB (courbe du haut) par rapport au Rapport Signal sur Bruit (RSB).

2.5.3 Influence de la structure d'observation

Ici nous considérons le cas I (défini dans la Section 2.3). La fonction d'utilité d'étape est fixée ici à w^{taux} (équation (2.55)). Les ensembles des puissances disponibles sont maintenant binaires : $\mathcal{X}_1 = \mathcal{X}_2 = \{0, P_{\max}\}$. Le changement important par rapport à la Section 2.5.2 est que nous ne considérons pas uniquement le cas d'observation parfaite. La Figure 2.9 représente les gains

$(\overline{Q}^{\star}_{X_1}(x_1), \overline{Q}^{\star}_{X_2}(x_2), \overline{\mathbf{Q}}^{\star}_{\mathbf{X_1X_2}}(\mathbf{x_1}, \mathbf{x_2}))$ in %	$x_1 = 0$	$x_1 = \frac{10}{3}$	$x_1 = \frac{20}{3}$	$x_1 = 10$
$x_2 = 0$	(44.4,50.4,**00.1**)	(02.6,50.4,**00.9**)	(08.0,50.4,**06.5**)	(45.0,50.4,**42.9**)
$x_2 = \frac{10}{3}$	(44.4,00.0,**00.0**)	(02.6,00.0,**00.0**)	(08.0,00.0,**00.0**)	(45.0,00.0,**00.0**)
$x_2 = \frac{20}{3}$	(44.4,00.0,**00.0**)	(02.6,00.0,**00.0**)	(08.0,00.0,**00.0**)	(45.0,00.0,**00.0**)
$x_2 = 10$	(44.4,49.6,**44.3**)	(02.6,49.6,**01.7**)	(08.0,49.6,**01.5**)	(45.0,49.6,**02.1**)

Tableau 2.7 – **Distributions marginales et distributions conjointes optimales** (exprimé en %) obtenus sous contraintes pour le cas somme des taux lorsque RSB = 10dB et qu'il y a quatre niveaux de puissance de transmission disponible $\{0, \frac{10}{3}, \frac{20}{3}, 10\}$.

$(\overline{Q}^{\star}_{X_1}(x_1), \overline{Q}^{\star}_{X_2}(x_2), \overline{\mathbf{Q}}^{\star}_{\mathbf{X_1X_2}}(\mathbf{x_1}, \mathbf{x_2}))$ in %	$x_1 = 0$	$x_1 = \frac{10}{3}$	$x_1 = \frac{20}{3}$	$x_1 = 10$
$x_2 = 0$	(47.5,47.5,**00.0**)	(00.0,47.5,**00.0**)	(00.0,47.5,**00.0**)	(52.5,47.5,**47.5**)
$x_2 = \frac{10}{3}$	(47.5,00.0,**00.0**)	(00.0,00.0,**00.0**)	(00.0,00.0,**00.0**)	(52.5,00.0,**00.0**)
$x_2 = \frac{20}{3}$	(47.5,00.0,**00.0**)	(00.0,00.0,**00.0**)	(00.0,00.0,**00.0**)	(52.5,00.0,**00.0**)
$x_2 = 10$	(47.5,52.5,**47.5**)	(00.0,52.5,**00.0**)	(00.0,52.5,**00.0**)	(52.5,52.5,**05.5**)

Tableau 2.8 – **Distributions marginales et distributions conjointes optimales** (exprimé en %) obtenus sans contraintes pour le cas somme des taux lorsque RSB = 10dB et qu'il y a quatre niveaux de puissance de transmission disponible $\{0, \frac{10}{3}, \frac{20}{3}, 10\}$.

relatifs par rapport à la stratégie CPS pour différentes structures d'observation. Les gains relatifs sont représentés comme une fonction du RSB.

- La courbe tout en haut (en pointillés) correspond au cas de communication sans coût et est donc une borne supérieure sur la performance ;
- La courbe juste en dessous correspond au cas d'observation parfaite ;
- La troisième correspond au cas d'observation imparfaite, et plus précisément au scénario de retour RSIB avec $e = 0.1$. Dans ce scénario, il est facile de vérifier que le RSIB peut prendre $m = 7$ valeurs distinctes ;
- La courbe la plus basse correspond également au cas d'observation imparfaite, avec une observation de type CBS et avec une probabilité d'erreur de 5% c'est-à-dire, $Z_1 \sim \mathcal{B}(0.05)$, $\mathrm{P}(Z_1 = 1) = 0.05$.

Toutes ces courbes montrent une tendance générale : il existe une grande variété de RSBs pour lesquels il y a un intérêt à concevoir une stratégie coordonnée de contrôle de puissance plus efficace que la stratégie CPS. En effet, pour les canaux d'observation avec des niveaux de bruits classiques, les gains potentiels en terme de performance obtenus grâce au *contrôle de puissance codé* sont assez robustes au bruit de communication. Spécifiquement, pour RSB= 10 dB, le gain relatif chute de 22% en observation parfaite à 18% avec la structure d'observation CBS. La figure nous permet aussi de quantifier la performance limite qu'un canal avec retour RSIB peut fournir en terme de coordination d'un réseau distribué à interférence, réseau composé de deux paires émetteurs/récepteurs ici. Des gains plus significatifs encore pourraient être exhibés en considérant une stratégie de contrôle de puissance de référence plus naïve telle que la stratégie de contrôle de puissance Nash, ou en considérant des scénarios avec plus d'interférence. Ces derniers peuvent être obtenus en augmentant le ratio $\frac{p_{ii}}{p_{ij}}, j \neq i$.

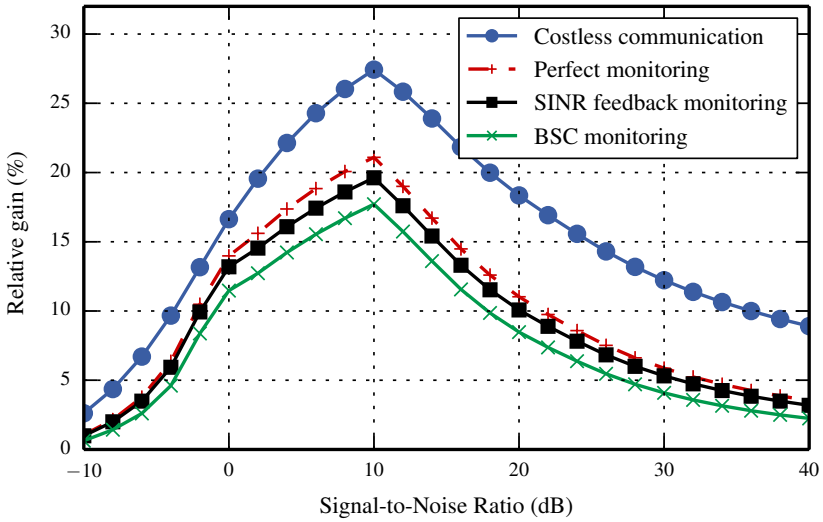

FIGURE 2.9 – **Influence de la structure d'observation** : gains relatifs en % en terme d'utilité espérée pour le cas somme des taux versus RSB (en dB). L'observation avec retour RSIB correspond à celle introduite en Figure 2.4. Le scénario de référence est CPS.

2.5.4 Influence de la connaissance sur l'état

Ici, nous considérons les mêmes choix de paramètres qu'en Section 2.5.3, mais cette fois l'observation est fixée à la structure d'observation CBS (toujours avec probabilité d'erreur égale à 0.05) et nous considérons à la fois le cas I et le cas II. Trois courbes sont représentées en Figure 2.10 :

- La courbe du haut représente toujours le cas de communication sans coût ;
- La courbe juste en dessous correspond au cas I ;
- La troisième et dernière courbe correspond au cas II, que nous obtenons en choisissant $|\mathcal{U}| = 10$.

Nous savons déjà que l'écart de performance entre les cas I et II est 0 en observation parfaite. Nous voyons que, pour une valeur classique du niveau de bruit du canal d'observation, le fait que l'Emetteur 2 ne connaisse pas les réalisations passées de l'information globale sur l'état du canal n'induit qu'une diminution des performances relativement faibles, et ce pour une cardinalité de l'alphabet de la variable aléatoire U raisonnablement petite.

2.5.5 Influence du schéma de coordination

Ici, pour rendre plus claires les interprétations, nous faisons des hypothèses simplificatrices. Nous considérons un cas particulier du canal à interférence, le **canal à accès multiple**. Il est composé de deux émetteurs et d'un récepteur. L'état du canal est maintenant composé de deux éléments (g_1, g_2) et plus quatre. Nous supposons que l'alphabet $x_0 = (g_1, g_2)$ est binaire :

$$\mathcal{X}_0 = \{x_0^0, x_0^1\} = \{(g_{\min}, g_{\max}), (g_{\max}, g_{\min})\} \tag{2.69}$$

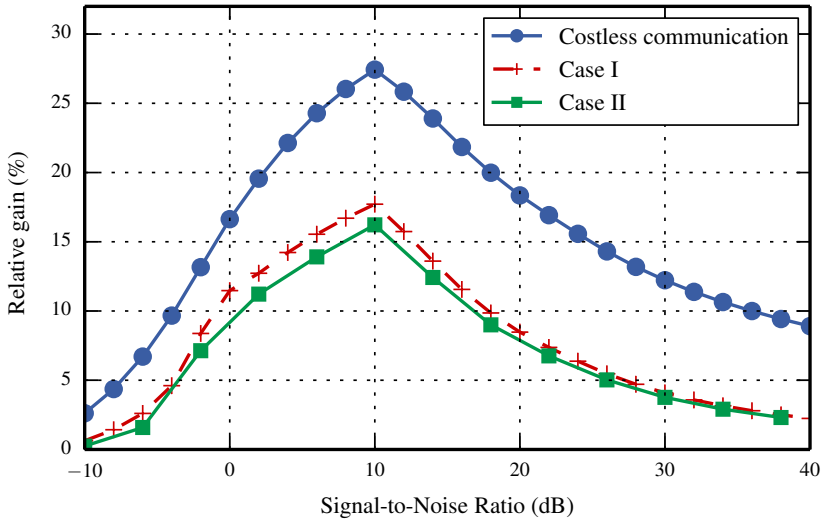

FIGURE 2.10 – **Comparaison des cas I et II** : gains relatifs en % en terme d'utilité espérée pour le cas somme des taux versus RSB (en dB). Le scénario de référence est CPS.

L'état du canal est une variable aléatoire de Bernoulli $X_0 \sim \mathcal{B}(p)$, avec $p = \mathrm{P}\big(X_0 = (g_{\min}, g_{\max})\big)$. Pour rendre la lecture du code de contrôle de puissance plus simple, nous utiliserons les notations "0" (respectivement "1") pour parler de la réalisation x_0^0 (respectivement x_0^1). La fonction d'utilité d'étape considérée est w^{RSIB}. Les différentes configurations peuvent être représentées sous forme de matrices, comme indiqué en Figure 2.11 : la matrice de gauche (respectivement de droite) correspond à l'état $x_0 = x_0^0 = (g_{\min}, g_{\max})$ (respectivement $x_0 = x_0^1 = (g_{\max}, g_{\min})$). Pour chaque matrice, le choix de la ligne (respectivement de la colonne) correspond à l'action de l'Emetteur 1 (respectivement 2). Les valeurs correspondantes de l'utilité sont données par les entrées de ces matrices.

	0	$P_{\max}$
0	0	20
$\mathbf{P}_{\max}$	1	$\simeq 10$

$x_0 = x_0^0$

	0	$P_{\max}$
0	0	1
$\mathbf{P}_{\max}$	20	$\simeq 10$

$x_0 = x_0^1$

FIGURE 2.11 – **Un jeu de contrôle de puissance**, donné sous forme matricielle pour le problème particulier, le canal à accès multiple. La fonction d'utilité w^{RSIB} est considérée, $\sigma^2 = 1$, $x_0 \in \{x_0^0 = (0.1, 2) \equiv \text{"0"}, x_0^1 = (2, 0.1)\} \equiv \text{"1"}$. Dans la suite, par soucis de simplicité, nous utilisons les notations $0 \equiv 0$, $1 \equiv P_{\max}$ pour le niveau de puissance de transmission des deux émetteurs.

La structure d'observation considérée est celle d'observation parfaite. Dans ce cadre, notre

but est d'évaluer l'influence du schéma de coordination et les performances du code de contrôle de puissance proposé. Ce dernier peut être vu comme un code source-canal *séparé* qui consiste en un code source avec distorsion et un code canal avec information adjacente. La longueur des mots de code est choisie à 3. L'encodeur source et le décodeur source sont donc définis par les fonctions :

$$\begin{array}{rccc} f_{\mathrm{S}}: & \mathcal{X}_0^3 & \to & \{m_0, m_1\} \\ & \underline{x}_0 & \mapsto & i \end{array}, \tag{2.70}$$

$$\begin{array}{rccc} g_{\mathrm{S}}: & \{m_0, m_1\} & \to & \mathcal{X}_2^3 \\ & i & \mapsto & \underline{x}_2 \end{array}. \tag{2.71}$$

Le code source choisi consiste à utiliser deux indices $\{m_0, m_1\}$ pour représenter les 8 séquences x_0^3 possibles. Le choix exact de f_{S} et g_{S} est donné après la description du code canal proposé. Pour définir l'encodeur canal clairement, nous utilisons l'indice de bloc b mais il faut noter que l'encodeur canal ne dépend pas de cet indice.
Sur chaque bloc b, l'encodeur canal implémente la fonction suivante :

$$\underline{x}_1^{(b)} = f_{\mathrm{C}}(\underline{x}_0^{(b)}, \underline{x}_2^{(b)}, i_{b+1}) \tag{2.72}$$

où $i_{b+1} = f_{\mathrm{S}}(\underline{x}_0^{(b+1)})$ est l'indice ou le message qui est associé à la séquence $x_2^{(b+1)}$. L'idée derrière la conception de l'encodeur canal f_{C} est la suivante. Si l'Emetteur 1 n'a pas à transmettre l'indice i_{b+1}, son encodage optimal sera d'exploiter sa connaissance de $(x_0^3(b), x_2^3(b))$ pour choisir la séquence $x_1^3(b)$ qui amène au paiement moyen maximal sur le bloc b. Cependant, pour communiquer l'indice i_{b+1}, l'Emetteur 1 devrait plutôt choisir de transmettre la séquence $x_1^3(b)$ ave le paiement moyen le plus grand sur le bloc b si $i_{b+1} = m_0$, ou la séquence $\underline{x}_1^{(b)}$ avec le second paiement moyen le plus haut sur le bloc b si $i_{b+1} = m_1$. Notons que l'Emetteur 2 est capable de décoder de manière parfaite cette façon d'encoder, en supposant sa connaissance de $\underline{x}_0^{(b)}$, $\underline{x}_2^{(b)}$, et $\underline{x}_1^{(b)}$ à la fin du bloc b. Formellement, f_{C} est définie comme suit : La séquence $\underline{x}_1$ est choisie comme

$$\underline{x}_1 = \underline{x}_1' \oplus \underline{d} \tag{2.73}$$

où l'addition modulo-2 est faite composante par composante.

$$\underline{x}_1' = \arg\max_{\underline{x}_1 \in \mathcal{X}_1^3} \sum_{t=1}^{3} w^{\mathrm{SINR}}(x_{0,t}, x_{1,t}, x_{2,t}), \tag{2.74}$$

$$\underline{d} = (0,0,0) \text{ if } i_{b+1} = m_0, \tag{2.75}$$

$$d^3 = \arg\max_{\underline{d} \text{ s.t. } \Omega(\underline{d})=1} \sum_{t=1}^{3} w^{\mathrm{SINR}}(x_{0,t}, x_{1,t}' \oplus d_t, x_{2,t}) \text{ if } i_{b+1} = m_1 \tag{2.76}$$

où Ω est la fonction de poids de Hamming, c'est-à-dire le nombre de 1 dans la séquence $\underline{d} \in \{0,1\}^3$. Si l'ensemble des argmax n'est pas un singleton, nous choisissons la séquence avec le plus petit poids de Hamming.
Pour compléter la construction, nous devons spécifier la conception du code source. Ici, nous choisissons les fonctions f_{S} et g_{S} qui maximise l'utilité moyenne $\mathbb{E}(w^{\mathrm{SINR}})$ en connaissant les opérations de codage canal. Le code source résultant d'une recherche exhaustive est donné en Tableau 2.13, et la code canal correspondant est donné en Tableau 2.12. L'expression détaillée de

$x_0^3(b)$	$x_2^3(b)$	i_{b+1}	$x_1^3(b)$
(0,0,0)	(1,1,1)	m_0	(0,0,0)
		m_1	(0,0,1)
(0,0,1)	(1,1,1)	m_0	(0,0,1)
		m_1	(0,0,0)
(0,1,0)	(1,1,1)	m_0	(0,1,0)
		m_1	(0,0,0)
(0,1,1)	(1,1,1)	m_0	(0,1,1)
		m_1	(0,0,1)
(1,0,0)	(1,1,1)	m_0	(1,0,0)
		m_1	(0,0,0)
(1,0,1)	(1,1,1)	m_0	(1,0,1)
		m_1	(0,0,1)
(1,1,0)	(0,0,1)	m_0	(1,1,0)
		m_1	(1,1,1)
(1,1,1)	(0,0,1)	m_0	(1,1,1)
		m_1	(1,1,0)

Tableau 2.12 – **Code canal proposé**, pour $p = \frac{1}{2}$.

x_0^3	Indice	x_2^3
(0,0,0)	m_0	(1,1,1)
(0,0,1)	m_0	(1,1,1)
(0,1,0)	m_0	(1,1,1)
(0,1,1)	m_0	(1,1,1)
(1,0,0)	m_0	(1,1,1)
(1,0,1)	m_0	(1,1,1)
(1,1,0)	m_1	(0,0,1)
(1,1,1)	m_1	(0,0,1)

Tableau 2.13 – **Codage et décodage source proposé**, pour $p = \frac{1}{2}$.

l'utilité espérée (dénotée $W_3(f_S, g_S \circ f_S)$) est fourni dans l'article de la section E.1.1 en Annexe E.

Les codes proposés peuvent être interprétés d'une manière très intuitive. Par exemple, la première ligne du Tableau 2.12 indique que l'Emetteur 1 ne transmet pas sur les trois étapes du bloc b si son canal est mauvais pour les trois étapes, c'est-à-dire pour $x_0^3 = (0,0,0)$. Il transmet uniquement s'il doit transmettre le message m_1. D'autre part, l'Emetteur 2 doit transmettre tout le temps sur le bloc considéré. Si nous regardons maintenant la dernière ligne du Tableau 2.12, nous voyons que l'Emetteur 1 est toujours on alors que l'Emetteur 2 est off les deux tiers du temps. Cela a l'air d'être sous-optimal puisque l'Emetteur 2 n'est pas toujours off. Cependant, puisque seulement deux séquences sont autorisées pour x_2^3, ces choix doivent aussi être bons pour les autres lignes. Il s'avère que cette séquence est en effet meilleure pour la ligne 7 du Tableau 2.12. En moyenne, nous pouvons vérifier que la fonction proposée est optimale pour le schéma de codage proposé.

Maintenant, comparons ce code court avec la meilleur performance possible que nous pourrions atteindre avec des codes longs. C'est le but de la Figure 2.14. Cette dernière représente w^{RSIB} comme une fonction du RSB en dB. La courbe du haut correspond au cas de communication sans coût, et la courbe juste en dessous représente la performance du meilleur schéma de contrôle de puissance codé qui est obtenue en résolvant le problème d'optimisation (2.40) pour le cas d'observation parfaite.
La troisième courbe en partant du haut (en pointillés) correspond à l'utilité espérée qui est atteint par le code de longueur 3 proposé. Nous voyons qu'il fonctionne très bien comparé aux deux courbes du dessus (les deux bornes supérieures que nous considérons). Cela offre des gains assez significatifs par rapport à la stratégie de référence CPS et cela fait clairement mieux que la stratégie de contrôle de puissance Nash.

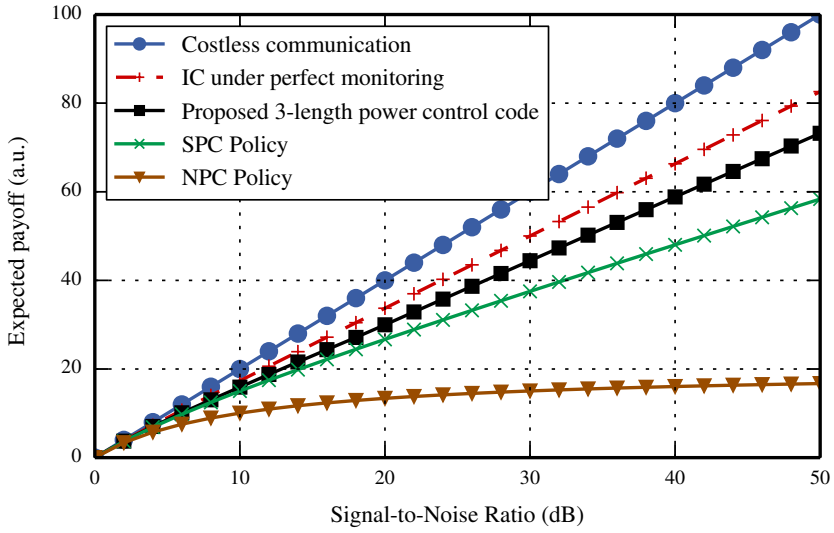

FIGURE 2.14 – **Utilité espérée vs RSB(dB) pour différents scénarios** : de la courbe du haut à celle du bas : communication sans coût, contrainte d'information dans le cas d'observation parfaite ; code de longueur 3 de contrôle de puissance ; stratégie de contrôle de puissance semi-coordonnée (CPS) ; stratégie de contrôle de puissance Nash (CPN).

2.6 Conclusion du chapitre

D'un point de vue pratique, un message clé de ce chapitre est qu'une stratégie de contrôle / d'allocation de ressource est un code, un code source-canal conjoint en général. Chaque agent du réseau distribué agit donc comme un encodeur qui essaye d'encoder sa connaissance sur l'état du réseau en une séquence d'actions. Les autres agents, qui ont une certaine observation de cette séquence jouent le rôle de décodeurs. Comme montré en Section 2.5.5, le but d'un bon code n'est pas de transmettre de l'information de manière fiable ou d'atteindre un niveau maximal de distorsion donné, mais plutôt de fournir une bonne utilité espérée. Pour cela, il doit implémenter un bon compromis entre envoyer de l'information (sur les réalisations futures de l'état) et atteindre une bonne valeur pour le paiement d'étape courant. De nouveaux codes qui atteignent ce compromis doivent donc être inventés. Compte tenu du grand nombre de fonctions d'utilité pour des problèmes d'allocation de ressource, une question pertinente est de savoir si des codes universels qui fonctionnent bien pour certaines classes d'utilité peuvent être construits.

De manière assez remarquable, puisqu'une stratégie de contrôle de puissance (distribué) est interprétée comme une code, la théorie de Shannon apparaît comme un outil naturel pour mesurer l'efficacité de ces stratégies. D'un point de vue théorique, il est essentiel de comprendre que la méthode utilisée pour trouver la meilleure performance en terme de coordination dans un réseau distribué est *générale*. Les hypothèses faites dans ce chapitre ont de grande chance d'être inadaptées pour certaines applications et les résultats proposés doivent être vus comme une première étape dans cette direction. Par exemple, comme mentionné en Section 2.1, une comparaison détaillée entre le contrôle de puissance codé et un algorithme itératif de type « water-filling » amènerait à considérer une structure d'information symétrique alors que nous étudions une structure d'information *asymétrique* dans ce chapitre. La méthodologie pour estimer les performances de bonnes stratégies de codage consiste à expliciter la bonne contrainte d'information en utilisant la théorie de l'information pour le problème de codage multi-source avec distorsion à travers des canaux multi-utilisateurs avec information adjacente et ensuite d'utiliser cette contrainte pour trouver l'utilité maximum (dans le cas d'intérêt commun) sous contrainte d'information ou l'ensemble des équilibres de Nash qui sont compatibles avec la contrainte (dans le cas d'intérêts différents, par exemple dans un jeu non-coopératif). Une observation clé dans ce chapitre est que la structure de communication d'un problème de décision à plusieurs agents correspond en fait à un canal multi-utilisateurs. La théorie de Shannon multi-terminal est donc non seulement pertinente pour des problèmes de communication mais aussi pour des problèmes de décisions multi-agents. Cette observation amène à de nouveaux problèmes ouvert pour les théoriciens de l'information puisque les problèmes de décisions définissent de nouveaux scénarios de communication.

Chapitre 3

Communication stratégique pour les smart grids

« Words can be meaningless. If they are used in such a way that no sharp conclusions can be drawn. »
– *The Meaning of It All : Thoughts of a Citizen-Scientist (2005)*, Richard P. Feynman.

Ce chapitre est consacré à un modèle de communication avec une asymétrie d'information comme au chapitre précédent. Cependant, nous supposons que les agents n'ont pas la même fonction d'utilité. Dans ce cas, il n'est plus possible d'étudier un problème d'optimisation et il faut caractériser l'ensemble des équilibres de Nash qui vérifie la contrainte d'information, ce qui est fait dans [46, Section 8.3] où un théorème Folk est donné, pour le cas d'observation parfaite de l'agent non informé des actions passées de l'agent informé. Dans ce chapitre nous supposons également l'observation parfaite, mais nous étudions un modèle différent où l'action de l'agent informé n'a pas d'influence (directe) sur le paiement. Dès lors, les actions de cet agent ont pour seul but de transmettre de l'information et nous sommes dans le cadre d'un jeu de « Cheap Talk », c'est-à-dire un jeu de communication où le message n'a pas d'influence directe sur l'utilité des agents. Par rapport au chapitre précédent, qui se focalisait sur l'étude de la coordination entre les deux agents qui ont un objectif commun, le but de ce chapitre est d'étudier l'aspect stratégique de la communication lorsque les agents ont chacun un objectif individuel. De plus, nous appliquons ce modèle à une problématique de Smart Grids [26], spécifiquement à la problématique de charge de véhicules électriques en fonction des contraintes du réseau d'électricité.

Nota Bene : toutes les preuves des résultats de ce chapitre peuvent être trouvées dans l'Annexe C.

26. Nous avons introduit les Smart Grids au Chapitre 1. Une autre référence peut être trouvée ici : http://www.smartgrids-cre.fr/index.php?p=definition-smart-grids.

3.1 Motivation et état de l'art

3.1.1 Introduction

Dans les réseaux électriques actuels, la production d'énergie est principalement commandée par la demande des consommateurs. Cependant, dans les Smart Grids il sera plus commun de voir le consommateur adapter sa consommation à la production, par exemple lorsqu'un mécanisme d'effacement [27] est implémenté ou lorsque la source d'énergie est solaire ou éolienne. Visiblement, le consommateur et l'agrégateur (c'est-à-dire l'entité qui prend les décisions concernant la demande du consommateur) auront des objectifs divergents de manière générale. En conséquence, il peut arriver que le consommateur reporte une demande plus élevée que son besoin réel en terme d'énergie dans le but de voir sa demande effective satisfaite. Comme, en pratique, la demande du consommateur (une usine, une maison, un propriétaire de véhicule électrique, etc.) en terme de besoin de puissance et la décision de la mesure dans laquelle délivrer cette puissance par l'agrégateur (une entreprise de service public, un opérateur de distribution de réseau, etc.) résultera souvent de procédures (automatiques) implémentées par des machines, un problème important d'ingénierie apparaît : comment concevoir un système de communication point-à-point où l'émetteur (le consommateur) et le récepteur (l'agrégateur) ont des objectifs divergents ?

En effet, le paradigme classique dans les systèmes de communication [79] suppose que l'émetteur et le récepteur ont un objectif commun (par exemple minimiser le taux de distorsion). Dans ce cas, trouver les stratégies des agents (les fonctions d'encodage et de décodage) peut être vu comme un problème de quantification scalaire, qui est lui même un cas particulier du problème de codage de source avec perte [26]. Lorsque les deux agents ont des intérêts divergents, le problème de codage (source/canal) doit être revu. En effet, même si la distribution de probabilité sur l'état aléatoire (le besoin du consommateur) est connue des deux agents, l'agent informé, sachant la divergence d'intérêts avec l'autre agent, ne maximisera pas son paiement moyen en révélant parfaitement la réalisation de l'état aléatoire. Il révélera plutôt une version dégradée de cette réalisation, et ce, même si le nombre de messages disponibles est infinie. Ce chapitre est une première étape dans le but de répondre à la question fondamentale posée ci-dessus. Nous faisons de plus un pont entre la littérature économique (notamment les jeux de « Cheap talk ») et celles du traitement du signal et des communications.

Spécifiquement, nous considérons deux agents (un consommateur, un agrégateur) qui discutent « hors ligne » à propos du schéma de signal qu'ils vont utiliser dans le but d'échanger de l'information de manière optimale. Ce schéma sera ensuite utilisé « en ligne » par les deux agents. Nous considérons donc un agrégateur qui a pour objectif de satisfaire la demande du consommateur mais également de minimiser le coût d'exploitation induit sur le réseau de distribution. Nous considérons particulièrement le vieillissement des transformateurs résidentiels ainsi que les pertes Joules comme coûts du réseau de distribution. D'autre part, l'objectif fondamental du consommateur est d'obtenir un niveau de puissance (ou d'énergie) le plus près possible de son vrai besoin. En prenant en compte un signal (message) reçu du consommateur, l'agrégateur décide finalement du niveau de puissance effectivement attribué au consommateur. Le message envoyé est supposé sans coût, c'est-à-dire qu'il n'a pas d'influence directe sur les utilités des deux agents. Ces utilités dépendent du besoin du consommateur ainsi que du niveau de puissance

27. Un mécanisme d'effacement consiste, en cas de déséquilibre entre l'offre et la demande d'électricité, à réduire provisoirement la consommation physique d'un site donné ou d'un groupe d'acteurs (par rapport à sa consommation « normale »). Cet effacement est déclenché de façon extérieure (les agents n'ont plus le pouvoir de décision).

effectivement attribué. Un des buts de ce travail est de construire un schéma de signalisation du consommateur vers l'agrégateur qui leur permettrait d'atteindre un consensus (un équilibre) sur comment communiquer en pratique (par rapport à une norme de communication adaptée). Il s'avère qu'en considérant un modèle simple mais néanmoins réaliste entre le consommateur et l'agrégateur, le problème à résoudre devient un jeu qui aura une formulation liée au problème de transmission stratégique de l'information [27] en économie, et sera également lié au problème de quantification.

3.1.2 État de l'art

En effet, le problème de transmission stratégique de l'information a été introduit dans [27] et développé en économie, comme par exemple l'interaction entre un recruteur et un demandeur d'emploi : ce dernier doit révéler plus ou moins d'information à propos de son « état », c'est-à-dire son niveau de qualification ou ses attentes en terme de salaire, alors que le recruteur doit décider ou non de lui donner le poste, en choisissant également le niveau de salaire accompagnant ce poste. Une étude récente de ses développements en économie peut être trouvée dans [82].
La transmission stratégique de l'information [27] est modélisée comme suit : un agent informé à propos de la réalisation d'un état aléatoire, envoie un message à un agent non informé, qui choisit une action en se basant sur le message reçu. Les utilités des deux agents ne dépendent que de la réalisation de l'état et de l'action prise par l'agent non informé, mais pas du message envoyé. Le message est donc gratuit au sens des utilités et nous parlons de jeu de « Cheap talk ». En toute généralité, les utilités des deux agents ne sont pas alignées, et un paramètre b (le biais) rend compte de la divergence sur les utilités des agents. Crawford et Sobel prouvent que les équilibres de Nash de ce jeu correspondent à une partition en intervalles de l'espace des états aléatoires pour l'agent informé, où à chaque intervalle est associé un message ; et une correspondance pour l'agent non informé, qui a chaque message reçu associe une action (différente). Tous ces équilibres sont des *équilibres partitionnels*. De plus, ils prouvent qu'il existe un entier qui dépend de b, tel que pour toutes les tailles de partition (ou de manière équivalente pour tous nombres de messages) de un à cet entier il existe un équilibre partitionnel. Il existe donc un nombre maximum de messages utilisés à l'équilibre.
La caractérisation des équilibres est en partie basée sur des hypothèses sur les fonctions d'utilités : elles doivent être concave en l'action de l'agent non informé, et elles doivent être croissante en l'action lorsque l'état est croissant. Cette condition traduit les préférences des agents, qui préfèrent une action plus grande pour un état plus grand. De plus, les équilibres sont caractérisés grâce à une condition d'arbitrage, qui traduit le fait que l'agent informé qui aurait observé un état à la frontière entre deux intervalles de sa partition, doit être indifférent entre envoyer le message de l'intervalle précédent et celui de l'intervalle suivant.

Ce modèle de transmission stratégique n'a pas encore été étudié dans des problèmes d'ingénierie, à quelques exceptions près [54] [68], qui ne considèrent pas l'application aux Smart Grids ou la connexion entre le codage et la quantification.

Deux extensions du modèle de Crawford et Sobel sont faites dans [45] : premièrement, l'auteur considère que les préférences des agents peuvent être alignées pour certains états. Dans ce cas, il peut ne plus exister de taille de partition maximale : pour tout entier naturel, un équilibre partitionnel existe. Deuxièmement, dans le modèle de [27], l'agent informé préfère toujours une action plus grande que l'agent non informé, quelque soit l'état. [45] généralise cela en considérant toutes les « directions » possibles entre les préférences des agents. Ces directions ont une influence sur la caractérisation des équilibres. Nous ferons des liens avec ce modèle dans les sections suivantes.

Une autre extension est de considérer des ensembles d'états, de messages, et d'actions vectoriels et non plus scalaires. Ceci est étudié dans [52] pour le cas où les deux agents ont la même fonction d'utilité. Les éléments des partitions d'équilibre ne sont plus des intervalles mais des cellules de Voronoï[28], et les différentes actions prises par l'agent non informé sont les centroïdes de ces ensembles. Les équilibres sont en fait trouvés en faisant de la quantification de l'espace continu des états.
L'extension vectorielle pose des problèmes difficiles, à commencer par le cas où les agents ont la même utilité, et bien plus lorsqu'il nous faut considérer deux utilités distinctes. En effet, les objets à manipuler sont plus compliqués. Nous faisons une discussion sur les dimensions supérieures à un en section 3.3.2.

Le reste de ce chapitre est organisé comme suit : en section 3.2, nous introduisons le modèle que nous étudions et montrons les différences avec les modèles existants. Un exemple introductif en section 3.2.3 permet de fixer les idées. Ensuite, nous expliquons les résultats théoriques importants liés à notre modèle en section 3.3. Enfin, nous illustrons ces résultats grâce à des simulations rendant compte de situations pratique en section 3.4.

3.2 Modèle de jeu de signaux et exemple introductif

3.2.1 Modèle de jeu de signaux

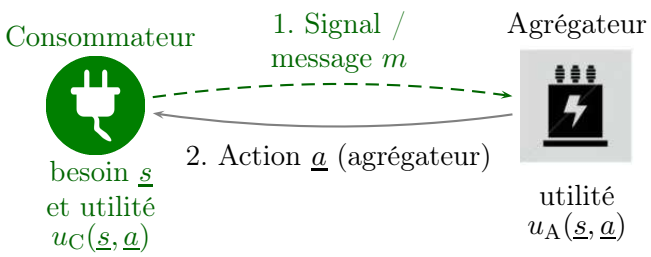

FIGURE 3.1 – **Schéma de signal pour un modèle de Smart Grids** : le consommateur a un vecteur de besoins en terme de puissance $\underline{s}$ qui est inconnu de l'agrégateur. Le consommateur envoie un message $m \in \{1, 2, \dots, M\}$ à l'agrégateur. Ce dernier choisit alors un vecteur d'actions $\underline{a}$, qui est le niveau de puissance effectivement alloué au consommateur. Le point clé est que l'agrégateur et le consommateur ont des fonctions d'utilités *non-alignées*.

Considérons le modèle illustré en Figure 3.1, que nous exposons en toute généralité en dimension T quelconque : deux agents, un consommateur et un agrégateur, interagissent pour savoir comment allouer de manière optimale les besoins (représentés par un vecteur de taille T, $T \in \mathbb{N}$) au consommateur. Le consommateur (typiquement un véhicule électrique) a un besoin de charge vectoriel $\underline{s} \in \mathcal{S} \subset [0,1]^T$, avec $\underline{s}_t \geq 0 \; \forall 1 \leq t \leq T$ le besoin à l'instant t, qui est représenté par une

28. Les cellules de Voronoï sont associées à un ensemble discret de points, appelés centroïdes. Une cellule de Voronoï est un ensemble de points de l'espace considéré qui sont les plus proches d'un certain centroïde plutôt que de tout autre centroïde.

variable aléatoire avec distribution de probabilité $p(\cdot)$ connue des deux agents. Le consommateur veut obtenir un vecteur de puissances allouées aussi proche que possible de son vrai besoin. L'agrégateur, qui choisira les puissances, ne connaît pas le vrai besoin du consommateur. Seule la distribution $p(\cdot)$ est connue. Un canal de communication est disponible du consommateur à l'agrégateur : le consommateur choisit un message $m \in \{1, 2, \dots, M\}$, $M < \infty$ (parmi un nombre fini donc), qui est observé de manière parfaite par l'agrégateur. Ce dernier choisira une action (un vecteur de niveaux de puissance alloués), dénotée $\underline{a}$, $\underline{a}_t \geq 0 \; \forall 1 \leq t \leq T$, en tenant compte du message qu'il a reçu. Sans perte de généralité, nous supposons que $(\underline{s}, \underline{a}) \in \mathcal{S}^2$. De plus, nous supposons que les fonctions d'utilités dépendent du besoin $\underline{s}$ et des niveaux de puissance alloués $\underline{a}$ mais pas (directement) des messages envoyés par le consommateur, c'est-à-dire que parler est « peu cher ». C'est pourquoi nous sommes dans le cadre d'un jeu de Cheap Talk comme dans [27].

Nous considérons les fonctions d'utilités suivantes, qui sont associées au consommateur et à l'agrégateur respectivement :

$$u_C(\underline{s}, \underline{a}) = -\|\underline{s} - \underline{a}\|_2^2 \,, \tag{3.1}$$

$$u_A(\underline{s}, \underline{a}) = -\|\underline{s} - \underline{a}\|_2^2 - bc(\underline{a}) \tag{3.2}$$

où $b \geq 0$ représente le poids mis sur le coût réseau c. Cela représente la pondération que nous imposons sur l'utilité du consommateur et le coût réseau. Faire varier le paramètre b permet de considérer différents points de vue selon la modélisation considérée (jeu d'équipe, jeu non-coopératif, etc). Dans la suite, nous suivrons la littérature des jeux de Cheap talk qui parle de *biais* entre les agents pour se référer au paramètre b.
Le consommateur se soucie de son propre besoin, alors que l'agrégateur se soucie à la fois de l'état du réseau (qui est représenté par la fonction de coût $c : \mathcal{S} \to \mathbb{R}$) et du besoin du consommateur. Le coût réseau peut représenter le facteur d'accélération du vieillissement d'un transformateur (résidentiel). Le vieillissement est accéléré de manière exponentielle lorsque le niveau de puissance est au dessus du niveau de charge nominal (voir [51]). Ce coût réseau peut également représenter les pertes Joules, qui sont le produit entre la résistance du câble et le courant au carré [34], qui d'un point de vue mathématique correspondent donc à une fonction quadratique. Dans ce chapitre, nous supposons que la fonction de coût c est croissante et convexe, et que les fonctions d'utilités sont connaissances communes.
Dans ce modèle, le consommateur veut à la fois atteindre son besoin en terme de puissance mais ne veut également pas dépasser ce niveau de puissance, ce qui pourrait entraîner un coût supplémentaire non nécessaire. Ce modèle peut être justifié également lorsque le besoin est interprété comme la quantité d'énergie désirée, par exemple pour recharger une batterie (voir [63]).

De plus, nous supposons que :

- une stratégie pour le consommateur est une fonction entre les réalisations de son besoin et l'ensemble fini des messages, qui peut être interprété comme une application

$$f : \left| \begin{array}{ccc} \mathcal{S} & \to & \{1, 2, ..., M\} \\ \underline{s} & \mapsto & m \end{array} \right. . \tag{3.3}$$

  Supposons de plus que la stratégie du consommateur appartient à une classe particulière de fonctions de la forme (3.3). Alors une stratégie est une partition de l'ensemble des états (du besoin), et nous pouvons représenter la stratégie du consommateur par un vecteur $\underline{S} =$

$(S_m)_{m\in\{1,\dots,M\}} \subset \mathcal{P}(\mathcal{S})^M$. Notons que lorsque $T = 1$, chaque élément S_m de la partition est simplement un intervalle : $S_m = [s_{m-1}, s_m]$, $\forall\, m \in \{1,\dots,M\}$, avec $s_{m-1} \leq s_m$;

- une stratégie pour l'agrégateur est donnée par une fonction entre l'ensemble des messages et l'ensemble des vecteurs d'action :

$$g : \left| \begin{array}{ccc} \{1,2,...,M\} & \to & \mathcal{S} \\ m & \mapsto & \underline{a} \end{array} \right. . \tag{3.4}$$

  De manière équivalente, nous pouvons représenter la stratégie de l'agrégateur par un vecteur de vecteurs (une matrice) $\underline{A} = (\underline{a}_m)_{m\in\{1,\dots,M\}} \in \mathcal{S}^M$.

Finalement, nous supposons que le consommateur maximise l'espérance, par rapport à la distribution de probabilité de son besoin, de sa fonction d'utilité $u_C(\cdot,\cdot)$:

$$W_C(f,g) = \int_{\underline{s}\in\mathcal{S}} u_C\Big(\underline{s}, g(f(\underline{s}))\Big)\, p(\underline{s})\, d\underline{s} \tag{3.5}$$

$$= -\int_{\underline{s}\in\mathcal{S}} \|\underline{s} - g(f(\underline{s}))\|_2^2\, p(\underline{s})\, d\underline{s} \tag{3.6}$$

et que l'agrégateur maximise l'espérance, par rapport à la distribution de probabilité du besoin, de sa fonction d'utilité $u_A(\cdot,\cdot)$:

$$W_A(f,g) = \int_{\underline{s}\in\mathcal{S}} u_A\Big(\underline{s}, g(f(\underline{s}))\Big)\, p(\underline{s})\, d\underline{s} \tag{3.7}$$

$$= -\int_{\underline{s}\in\mathcal{S}} \Big[\|\underline{s} - g(f(\underline{s}))\|_2^2 + bc\Big(g(f(\underline{s}))\Big)\Big]\, p(\underline{s})\, d\underline{s} \tag{3.8}$$

Lorsque nous supposons que les stratégies sont représentées par des matrices $\underline{S}$ et $\underline{A}$, les utilités sont des fonctions de ces quantités et s'écrivent comme suit :

$$U_C(\underline{S},\underline{A}) = -\sum_{m=1}^{M} \int_{\underline{s}\in S_m} \|\underline{s} - \underline{a}_m\|_2^2\, p(\underline{s})\, d\underline{s} \tag{3.9}$$

and

$$U_A(\underline{S},\underline{A}) = -\sum_{m=1}^{M} \int_{\underline{s}\in S_m} \Big[\|\underline{s} - \underline{a}_m\|_2^2 + bc(\underline{a}_m)\Big]\, p(\underline{s})\, d\underline{s} \tag{3.10}$$

3.2.2 Discussion

Le problème d'optimisation des fonctions f et g est en fait lié à la quantification. Si f et g sont optimisées séparément, le problème peut être interprété comme un jeu où un agent choisit f et l'autre choisit g. Deux points très importants :

- La *différence fondamentale* entre la quantification pour une unique fonction d'utilité et le cas (comme dans le modèle de [27]) où les valeurs des utilités des deux agents sont distinctes pour toute réalisation de l'état aléatoire est la suivante : toute les ressources de communication ne sont pas forcément utilisées. En effet, en quantification, quelque soit le nombre de messages disponibles, ne pas utiliser tous ces messages est sous optimal pour

maximiser la fonction : plus l'information est précise, plus la fonction d'utilité est maximisée. En revanche, dans le modèle avec utilités distinctes, il existe un nombre maximum de messages utilisés à l'équilibre [27, Théorème 1]. Si les ressources de communication permettent d'utiliser plus de messages, ces messages supplémentaires seront inutiles pour une stratégie d'équilibre.

- Si les utilités sont communes (par exemple lorsque $b = 0$) et le nombre de messages M est fixé et fini, cela défini un jeu de potentiel [71][76]. Pour ce type de jeu, il est connu que la procédure itérative consistant à optimiser la fonction d'utilité par rapport à f pour g fixée, puis l'optimiser par rapport à g pour la nouvelle fonction f, et ainsi de suite, converge vers un équilibre de Nash. Cette procédure est appelée la dynamique de meilleure réponse séquentielle en théorie des jeux. L'algorithme de Lloyd-Max (pour une unique fonction d'utilité donc, et en dimension $T = 1$) implémente précisément cette procédure et converge vers un équilibre de Nash. En effet, les points d'intersection des meilleures réponses sont précisément les équilibres du jeu. Dans ce chapitre, nous étudions cette dynamique et les meilleures réponses dans le cas de fonctions d'utilités différentes.

Il est maintenant important de préciser les liens avec les modèles existants. Les différences avec le modèle de Crawford et Sobel [27] sont tout d'abord que nous considérons un besoin et des actions vectoriels alors que [27] considère uniquement un modèle scalaire. D'autre part, nous considérons des utilités bayésiennes pour les deux agents, alors que dans le modèle originel l'agent informé maximise sa fonction d'utilité pour chaque réalisation fixée de son état.
De plus, une hypothèse clé dans ce modèle *scalaire* est que les utilités (instantanées) vérifient les conditions suivantes :

$$\forall\, i \in \{\mathrm{A}, \mathrm{C}\}\,, \begin{cases} \forall a, & \exists s, \frac{\partial u_i}{\partial a}(s,a) = 0 \\ \forall (a,s), & \frac{\partial^2 u_i}{\partial a^2}(s,a) < 0 \\ \forall (a,s), & \frac{\partial^2 u_i}{\partial s \partial a}(s,a) > 0 \end{cases} . \tag{3.11}$$

La dernière hypothèse est appelée la condition de « single crossing ». Ces hypothèses garantissent que les utilités ont un unique maximum, et également que le consommateur préfère une action plus grande que l'agrégateur pour chaque réalisation de l'état aléatoire.

Lorsque nous prenons $T = 1$ dans notre modèle (donc que nous passons au cas scalaire), les fonctions d'utilités (3.1) et (3.2) que nous avons choisies vérifient bien les hypothèses du modèle de [27].

Dans le cas général où $T \geq 1$, nous sommes dans le cadre de [52] mais avec des utilités *non-alignées* ce qui est en fait une généralisation de la quantification vectorielle au cas où deux fonctions doivent être prise en compte.

3.2.3 Un exemple illustratif

Nous considérons dans cette section un exemple simple pour $T = 1$ (un cas scalaire donc), qui est utile pour expliquer les liens entre le modèle proposé et celui introduit dans [27]. Soit p la distribution de probabilité uniforme sur $\mathcal{S} = [0, 1]$. Une stratégie pour le consommateur est une partition $(S_m)_{m\in\{1,\dots,M\}}$ où $S_m = [s_{m-1}, s_m]$, $s_0 = 0$ et $s_M = 1$. De manière équivalente, nous pouvons voir une stratégie du consommateur comme un vecteur $(s_m)_{m\in\{0,\dots,M\}} = (0, s_1, \dots, s_{M-1}, 1)$ que nous supposons ordonné sans perte de généralité. Une stratégie de l'agrégateur est un vecteur $(a_m)_{m\in\{1,\dots,M\}}$ où chaque a_m, $m \in \{1, \dots, M\}$ est associé à l'intervalle $S_m = [s_{m-1}, s_m]$.

Fixons $\epsilon > 0$ et $b \geq 0$. Pour cet exemple, nous considérons les fonctions d'utilités suivantes :

$$U_C^{\epsilon}((s_m)_m, (a_m)_m) = -\sum_{m=1}^{M} \int_{s_{m-1}}^{s_m} (a_m - (s+\epsilon))^2 \, p(s) \, ds \tag{3.12}$$

$$U_A((s_m)_m, (a_m)_m) = -\sum_{m=1}^{M} \int_{s_{m-1}}^{s_m} \left[(a_m - s)^2 + b a_m^2\right] \, p(s) \, ds \tag{3.13}$$

Les fonctions $a \longmapsto -(a-(s+\epsilon))^2$ et $a \longmapsto -(a-s)^2 - ba^2$ sont convexes, vérifient la condition de single crossing (équation (3.11)) et les préférences des agents sont strictement différentes pour toutes réalisations du besoin s :

$$\arg\max_a u_C^{\epsilon}(a, s) = s + \epsilon > \arg\max_a u_A(a, s) = \frac{s}{1+b}. \tag{3.14}$$

Alors, pour un nombre de messages M fixé, un équilibre existe et peut être explicitement caractérisé grâce à [27, Théorème 1]. Nous commençons par le calcul de la meilleure réponse de l'agrégateur, en prenant en compte la stratégie du consommateur comme un vecteur fixé $(s_m)_{m\in\{0,\dots,M\}}$. Pour chaque intervalle $[s_{m-1}, s_m], m \in \{1, \dots, M\}$, la meilleure réponse de l'agrégateur est :

$$a_m^* \triangleq \arg\max_a -\int_{s_{m-1}}^{s_m} \left[(a-s)^2 + ba^2\right] p(s) \, ds \tag{3.15}$$

$$= \frac{\int_{s_{m-1}}^{s_m} s p(s) \mathrm{d}s}{(1+b)\int_{s_{m-1}}^{s_m} p(s) \, \mathrm{d}s} \tag{3.16}$$

$$= \frac{s_{m-1} + s_m}{2(1+b)} \ , \tag{3.17}$$

où la dernière égalité est vraie grâce à l'hypothèse sur $p(\cdot)$ de distribution uniforme.

Pour le consommateur, nous devons trouver les valeurs optimales s_m^*, $m \in \{1, \dots, M-1\}$ (s_0 et s_M sont fixés à 0 et 1 respectivement), donc qui maximisent U_C^{ϵ} en prenant en compte la stratégie $(a_m)_m$ de l'agrégateur. Les valeurs sont, $\forall\, m \in \{1, \dots, M-1\}$:

$$(s_m^*)_m = \arg\max_{(s_m)_m} U_C^{\epsilon}((s_m)_m, (a_m)_m) \ . \tag{3.18}$$

Comme nous cherchons un point fixe de la fonction composée des meilleures réponses, nous allons trouver un vecteur optimal $(s_m^*)_{m\in\{0,\dots,M\}}$, en prenant en compte la meilleure réponse de l'agrégateur a_m^* trouvée en (3.17) :

$$(s_m^*)_m = \arg\max_{(s_m)_m} U_C^{\epsilon}((s_m)_m, (a_m^*)_m). \tag{3.19}$$

La condition du premier ordre pour la maximisation du problème (3.19) est, $\forall\, m \in \{1, \dots, M-1\}$:

$$\left(\frac{s_{m-1} + s_m}{2(1+b)} - (s_m + \epsilon)\right)^2 = \left(\frac{s_m + s_{m+1}}{2(1+b)} - (s_m + \epsilon)\right)^2 \ . \tag{3.20}$$

Notons que cette équation est exactement la *condition d'arbitrage* caractérisée dans [27]. Cette condition peut être expliquée intuitivement : si le besoin du consommateur s est exactement

s_m pour un certain m, le consommateur doit être indifférent (c'est-à-dire doit obtenir le même paiement) entre envoyer le message de l'intervalle précédent S_m ou celui de l'intervalle suivant S_{m+1}.
L'équation (3.20) implique :

$$s_{m+1} = s_{m-1}, \tag{3.21}$$

ce qui est impossible par définition de la partition $\mathbf{s}$, ou :

$$s_{m+1} = (2+4b)s_m - s_{m-1} + 4(1+b)\epsilon. \tag{3.22}$$

Cette équation de récurrence a la solution suivante :

$$\forall\, m,\ s_m^\epsilon = k_+\alpha_+^m + k_-\alpha_-^m - \frac{1+b}{b}\epsilon\ , \tag{3.23}$$

avec

$$\alpha_\pm \triangleq 1 + 2b \pm 2\sqrt{b^2+b}\ , \tag{3.24}$$

et

$$\begin{cases} k_+ \triangleq \frac{s_1^\epsilon + \frac{1+b}{b}\epsilon - \frac{1+b}{b}\epsilon\alpha_-}{4\sqrt{b^2+b}} \\ k_- \triangleq \frac{\frac{1+b}{b}\epsilon\alpha_+ - s_1^\epsilon - \frac{1+b}{b}\epsilon}{4\sqrt{b^2+b}} \end{cases}\ . \tag{3.25}$$

Observons que cette solution est paramétrée par la première valeur s_1^ϵ. Cette valeur doit être fixée de manière à obtenir $s_M^\epsilon = 1$, ce qui fournit une partition de taille M.
Notons également que lorsque $\epsilon \to 0$, cette solution converge vers $\mathbf{s}^0 = (s_m^0)_m$ avec :

$$\forall m, s_m^0 = \frac{\alpha_+^m - \alpha_-^m}{\alpha_+^M - \alpha_-^M}\ , \tag{3.26}$$

qui est un équilibre du jeu défini précédemment, pour M fixé. En particulier, pour tout $M \in \mathbb{N}$ il existe une partition de taille M. Ce qui est cohérent et cela aurait pu être établi à priori en utilisant [45, Théorème 4]. En effet, le biais b de notre exemple est « vers l'extérieur », c'est-à-dire que l'action préférée de l'agent informé (le consommateur) est plus extrême que celle de l'agent non informé (l'agrégateur) pour des états extrêmes. Dans notre exemple, cela s'écrit :

$$\left[\arg\max_a u_A(a,0), \arg\max_a u_A(a,1)\right] = \left[0, \frac{1}{1+b}\right] \tag{3.27}$$

$$\subseteq \left[\arg\max_a u_C^{\epsilon=0}(a,0), \arg\max_a u_C^{\epsilon=0}(a,1)\right] \tag{3.28}$$

$$= [0,1]\ . \tag{3.29}$$

Lorsque $\epsilon > 0$, les conditions (3.11) s'appliquent et il existe une taille maximale de partitions d'équilibre comme le montre le tableau 3.2.

Nous revenons maintenant au cas avec $\epsilon = 0$, c'est-à-dire celui qui correspond aux utilités de la forme (3.1) et (3.2). Il est intéressant de noter que, pour une taille donnée M, une partition d'équilibre varie avec b :

$b \diagdown \epsilon$	10^{-7}	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}
10^{-3}	157	120	84	49	21	7	2
10^{-2}	60	50	38	27	15	7	2
10^{-1}	22	19	16	12	8	5	2
0.5	11	10	9	7	5	3	2

Tableau 3.2 – **Taille maximale des partitions d'équilibre** : ce tableau indique la taille maximale des partitions d'équilibre en fonction de ϵ et du poids sur le coût réseau b. Nous observons que la taille est décroissante lorsque ϵ croît pour b fixé et décroissante lorsque b croît pour ϵ fixé.

- pour b petit, nous avons :

$$\alpha_{\pm} = 1 \pm 2\sqrt{b} + o(\sqrt{b}). \tag{3.30}$$

  Donc :

$$\forall m,\ 1 \leq m \leq M-1,\ s_m = \frac{4m\sqrt{b}}{4M\sqrt{b}} = \frac{m}{M} + o(\sqrt{b})\ , \tag{3.31}$$

  et la partition est uniforme.

- pour b grand :

$$\begin{cases} \alpha_+ = 4b + o(b) \\ \alpha_- = o(1) \end{cases} \tag{3.32}$$

  ce qui implique

$$\forall m,\ 1 \leq m \leq M-1,\ s_m = \frac{(4b)^m}{(4b)^M} = (4b)^{m-M} + o(b)\ , \tag{3.33}$$

  et la partition d'équilibre est « proche » d'une partition avec un intervalle unique $[0,1]$ puisque tous les s_m sont concentrés en 0.

Pour illustrer tout cela, nous avons tracé les partitions de taille 5 et les actions optimales correspondantes pour différentes valeurs du paramètre b en Figure 3.3. Nous voyons que pour b petit, le problème se rapproche du problème de quantification et le consommateur donne le maximum d'information possible à l'agrégateur. De plus, comme la probabilité de l'état est uniforme, lorsque b tend vers 0, les actions ont tendance à être au centre des intervalles. b grand implique que l'agrégateur ne choisira pas de grandes valeurs pour ses actions, c'est pourquoi les points extrêmes des intervalles sont plus proches de zéro : donner de l'information pour un "grand" état n'aura pas d'intérêt.

De plus, pour une valeur fixée du paramètre b et toujours pour un nombre de messages $M = 5$, nous montrons l'influence de la distribution de probabilité sur l'état en figure 3.4. Le fait que les intervalles soient « concentrés » autour d'une valeur pour une loi de nature gaussienne s'explique par le fait que la probabilité que l'état soit dans cette partie est plus grande donc donner de l'information à cet endroit amène à un gain plus important.

Cet exemple illustre bien les différents résultats que nous pouvons obtenir en fonction des hypothèses faites sur le modèle étudié. Dans la section suivante, nous détaillons les résultats théoriques de ce modèle.

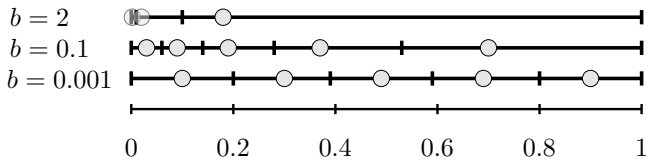

FIGURE 3.3 – **Partitions optimales de taille 5** : pour $M = 5$, nous avons représenté les partitions optimales pour les valeurs du paramètre b suivantes : $b \in \{0.001, 0.1, 2\}$. La distribution de probabilité prise sur l'état (le besoin) est la probabilité uniforme.

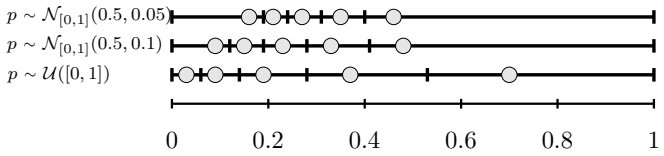

FIGURE 3.4 – **Partitions optimales en fonction de la probabilité sur l'état** : pour un nombre de messages $M = 5$, un biais $b = 0.1$, nous faisons varier la distribution de probabilité $p(\cdot)$: nous prenons une loi normale avec deux variances différentes (0.05 et 0.1) ainsi que la probabilité uniforme sur $[0, 1]$.

3.3 Résultats théoriques

3.3.1 Le cas de la dimension 1

Nous détaillons ici les résultats théoriques que nous avons dans le cas scalaire, c'est-à-dire pour $T = 1$. Premièrement, nous commençons par énoncer un résultat d'existence d'équilibre de Nash mixte dans notre modèle, pour un nombre de messages M fixé.

Proposition 3.1 (Existence). *Pour tout M fixé, il existe un équilibre de Nash mixte avec M messages.*

Ce résultat d'existence ne garantit pas qu'un équilibre partitionnel de n'importe quelle taille existe. En effet, il peut y avoir des cellules de la partition qui dégénèrent, c'est-à-dire dont le poids tend vers 0, ce qui implique que les messages associés à ces cellules sont envoyés avec probabilité nulle. Le nombre de cellules (messages/actions) peut donc être borné.

Si nous supposons de plus que les fonctions d'utilités des deux agents sont concaves, alors les équilibres de Nash seront purs comme dans [27], qui donne une méthode constructive pour les expliciter.

Nous exhibons de plus une condition suffisante d'unicité de l'équilibre, dans le cas où la fonction de coût $c(\cdot)$ est quadratique. Cette condition sera vérifiée dans des cas pratiques d'un intérêt important pour les réseaux de véhicules électriques. Elle est de plus moins restrictive que celle de [27], qui suppose également que les fonctions d'utilités ne dépendent que de la différence $s - a$ entre le besoin et l'action.

Proposition 3.2 (Condition suffisante d'unicité de l'équilibre). *Supposons que la fonction de coût c est quadratique, c'est-à-dire : $c(a) = \|a\|_2^2 = a^2$. Définissons*

$$p_{max} = \arg\max_{s \in [0,1]} p(s), \tag{3.34}$$

et

$$p_{min} = \arg \min_{s \in [0,1]} p(s) \tag{3.35}$$

Si

$$\frac{p_{max}}{p_{min}} \leq \sqrt{1+b} \tag{3.36}$$

et le nombre de messages M est fixé, alors l'équilibre de Nash de taille M est unique.

Maintenant que nous avons montré l'existence et l'unicité (dans certains cas) de l'équilibre de Nash, une question fondamentale est de trouver une méthode efficace pour les calculer. Nous analysons donc dans la suite la dynamique de meilleure réponse pour un nombre de messages M fixé, et nous montrons qu'elle converge si une hypothèse (raisonnable) est faite sur la distribution de probabilité du besoin du consommateur.

Pour commencer, nous devons prouver quelques propriétés utiles sur la dynamique dans notre contexte. Premièrement, il est possible d'exhiber une forme explicite des meilleures réponses des deux agents (toujours dans le cas de la dimension un). C'est ce que caractérise la proposition suivante :

Proposition 3.3 (Caractérisation explicite des meilleures réponses)**.** *Supposons que $p(\cdot) > 0$, que le nombre de messages M est fixé, et que la fonction de coût c est convexe et strictement croissante. Alors les meilleures réponses du consommateur et de l'agrégateur sont respectivement :*

$$BR^C(\mathbf{a}) = (s_m)_{m \in \{0,\dots,M\}} \tag{3.37}$$

$$\triangleq \begin{cases} 0 \ si \ m = 0 \\ \frac{a_m + a_{m+1}}{2} \ si \ 1 \leq m \leq M-1 \\ 1 \ si \ m = M \end{cases}, \tag{3.38}$$

et

$$BR^A(\mathbf{s}) \triangleq \left(\max(0, h^{-1}(2\frac{\int_{s_{m-1}}^{s_m} vp(v)dv}{\int_{s_{m-1}}^{s_m} p(v)dv})) \right)_{m \in \{1,\dots,M\}} \tag{3.39}$$

où nous définissons la fonction : $h : a \longmapsto 2a + bc'(a)$.

Notons que la meilleure réponse du consommateur est exactement une partition de Voronoï. Observons également que dans le cas où la fonction de coût est $c(a) = a^2$, la meilleure réponse de l'agrégateur se réduit à la forme

$$BR^A(\mathbf{s}) = \left(\frac{\int_{s_{m-1}}^{s_m} sp(s)\mathrm{d}s}{(1+b)\int_{s_{m-1}}^{s_m} p(s)\mathrm{d}s} \right)_{m \in \{1,\dots,M\}} \tag{3.40}$$

qui devient, quand $b = 0$:

$$BR^A(\mathbf{s}) = \left(\frac{\int_{s_{m-1}}^{s_m} sp(s)\mathrm{d}s}{\int_{s_{m-1}}^{s_m} p(s)\mathrm{d}s} \right)_{m \in \{1,\dots,M\}} \tag{3.41}$$

qui est exactement ce que nous trouvons en quantification.

Nous regardons maintenant la notion de dégénérescence des cellules [36] : une cellule dégénère si ses points extrêmes sont confondus. Dans le cas de la dimension un, cela veut dire que, pour $m \in \{0, 1, \dots, M\}$ si $s_{m-1} = s_m$ alors la cellule $[s_{m-1}, s_m]$ est dégénérée. C'est une notion très importante pour que la dynamique de meilleure réponse soit bien définie à chaque itération et que le calcul des intégrales soit possible. Les auteurs de [36] font un raisonnement similaire dans le cas de la quantification et trouve une borne uniforme sur la taille des cellules.

En utilisant la caractérisation des meilleures réponses, nous pouvons prouver maintenant une propriété fondamentale de la dynamique de meilleure réponse, concernant l'ordre de dégénérescence des cellules de la partition utilisée par le consommateur à l'équilibre. La proposition suivante indique que, si des cellules dégénèrent, alors ces cellules doivent forcément être celles les plus proches de zéro :

Proposition 3.4 (Ordre de dégénérescence de la partition du consommateur)**.** *Supposons que $p(\cdot) > 0$ et que la fonction de coût c est convexe. Alors, si des cellules de la partition dégénèrent pendant la dynamique de meilleure réponse, c'est-à-dire si $s_{m-1}^{(n)} = s_m^{(n)}$ pour une certaine étape n du processus itératif, alors elles dégénèrent dans l'ordre croissant en partant de la cellule la plus proche de zéro.*

Notons ici que nous prouvons que si des cellules dégénèrent, elles le font en zéro. Ceci est dû au cas particulier de fonctions d'utilité choisies dans ce chapitre, et le fait que le biais b est choisi strictement plus grand que zéro. Pour d'autres fonctions d'utilités, il peut arriver que les points extrêmes des intervalles d'une partition d'équilibre se regroupent en un, ou à un autre point entre zéro et un. Ceci est en fait lié au sens du biais comme défini dans [45] : le sens du biais nous permet de savoir où la dégénérescence des cellules va se produire.

Nous pouvons maintenant énoncer une condition suffisante pour avoir des cellules de la partition du consommateur qui ne dégénèrent jamais au cours de la dynamique de meilleures réponses, c'est-à-dire $\forall n$, $s_1^{(n)} > 0$.

Proposition 3.5 (Condition suffisante de non-dégénérescence des partitions pendant la dynamique de meilleure réponse)**.** *Fixons $M \in \mathbb{N}$. Supposons que $p(\cdot) > 0$ et que la fonction de coût c est convexe. Si $c'(0) = 0$, alors les partitions d'équilibre ne dégénèrent pas tout au long de la dynamique de meilleure réponse.*

Lorsque cette condition suffisante sera vérifiée, la dynamique de meilleure réponse sera bien définie à chaque itération. Ce qui est essentiel pour le théorème suivant sur la convergence de la dynamique. Pour cela nous avons besoin d'un lemme, qui montre qu'un résultat proposé par [86], peut être étendu au cadre étudié ici :

Lemme 3.6 (Wu [86])**.** *Fixons le nombre M de messages et considérons les fonctions d'utilités (3.9) et (3.10). Si les actions de l'agrégateur (respectivement les points extrêmes des intervalles de la partition du consommateur) sont croissantes comme une fonction des points extrêmes de l'intervalle auquel il est associé (respectivement des actions), alors l'algorithme de Lloyd-Max converge.*

La preuve est la même que celle proposée dans [86], qui considère le cas où les agents maximisent la même fonction d'utilité (pour nous, cela correspond au cas $b = 0$). La fonction d'utilité

doit appartenir à une classe de fonctions qui dépend de la réalisation de l'état s, de la différence entre cette réalisation et l'action a, et doit être symétrique en $s-a$. Si $b=0$, notre fonction vérifie bien cela. Nous étendons cela au cas général $b \geq 0$ en utilisant la preuve pour deux fonctions d'utilités distinctes.

Nous avons finalement le résultat suivant sur la convergence de la dynamique de meilleure réponse :

Théorème 3.7 (Convergence de la dynamique de meilleure réponse)**.** *Supposons que $p(\cdot)$ est une densité de probabilité continue et strictement positive sur $[0,1]$, et que la dynamique de meilleure réponse est bien définie à chaque itération. Alors cette dynamique converge.*

Il est très important pour l'application pratique d'avoir la convergence de la dynamique de meilleure réponse, comme nous le verrons en section 3.4.

3.3.2 Discussion en dimension supérieure

Pour étudier ce modèle en dimension quelconque $T \geq 1$, de nouveaux problèmes se posent. Pour commencer, les objets à manipuler sont bien plus complexes, notamment les stratégies (et les espaces dans lesquels elles vivent) qui sont des vecteurs de vecteurs (voir par exemple les équations (3.9) et (3.10)). Les théorèmes d'existence d'équilibres de Nash ainsi que ceux de convergence de la dynamique de meilleure réponse ne s'étendent pas de manière triviale.
De plus, [52] est un des seuls articles qui traite de la dimension T quelconque, et ce pour une unique fonction d'utilité, donc un cas particulier de jeu. Le fait que les utilités soient les mêmes pour les deux agents permet d'avoir un jeu de potentiel et d'utiliser les propriétés de ce jeu, ce qui n'est plus possible lorsque les deux agents ont chacun leur utilité propre.
Enfin, d'un point de vue numérique, trouver les cellules de la partition (qui sont des polytopes convexes) n'est pas un problème compliqué [29], en revanche calculer leur volume pour pouvoir connaître la valeur de l'utilité des agents est un problème difficile [53] [59] . Il existe cependant des algorithmes pratiques se basant sur des méthodes de Monte-Carlo pour estimer avec une précision arbitrairement faible ces volumes [35] [41]. Ces méthodes pratiques convergent de plus en temps polynomial. Nous les utilisons pour notre application numérique lorsque $T > 1$.

3.4 Application numérique

Dans cette section, nous développons une application numérique de nos résultats, où nous fixons la fonction de coût à

$$c(a) = ||a||_2^2 \tag{3.42}$$

ce qui correspond aux pertes Joules. Nous supposons de plus que le besoin en terme de puissance du consommateur suit une loi *uniforme* sur $\mathcal{S}$, et nous étudions l'influence de différents paramètres de notre problème.
Dans la suite, nous considérons que $T=2$, c'est-à-dire que le consommateur a deux intervalles de temps et un besoin spécifique pour chaque intervalle.

29. Il est par exemple possible d'utiliser la Toolbox Matlab MPT (http://people.ee.ethz.ch/~mpt/3/) qui implémente cela.

3.4.1 Influence du biais b

Tout d'abord, nous illustrons le fait que même si un équilibre mixte existe pour tout nombre de messages M, il peut arriver que des cellules de la partition d'équilibre dégénèrent et que seulement un nombre fini borné de messages soit utilisé à l'équilibre. En Figure 3.5, nous représentons pour le cas scalaire $T = 1$ le nombre de cellules qui ne dégénèrent pas à l'équilibre, c'est-à-dire le nombre de cellules qui ont un volume supérieur à 1% du volume total de la partition. Nous voyons que plus b est grand, plus le nombre de cellules maximal qui ne dégénèrent pas est petit, ce qui nous permet de retrouver le résultat de [27] sur la taille maximale d'une partition d'équilibre. Lorsque b tend vers 0, cette taille maximale augmente et la partition d'équilibre tend vers une discrétisation très fine de l'ensemble des besoins du consommateur.

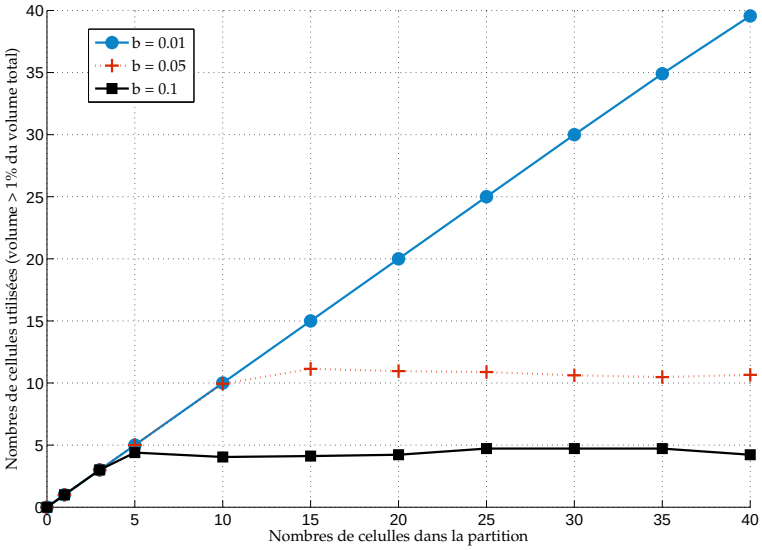

FIGURE 3.5 – **Nombre de cellules utilisées** : pour trois valeurs du paramètre b différentes ($b \in \{0.01, 0.05, 0.1\}$), nous représentons le nombre de cellules qui ne dégénèrent pas à l'équilibre en fonction du nombre de cellules total dans la partition d'équilibre. Une cellule ne dégénère pas si son volume est supérieur à 1% du volume total de la partition.

Deuxièmement, nous comparons les utilités espérées du consommateur et de l'agrégateur par rapport au cas sans communication, qui correspond à une partition d'équilibre avec un seul élément, et donc un seul message, ce qui ne donne pas d'information. Cette référence est très importante car l'équilibre sans communication [30] existe toujours et correspond au minimum que les agents peuvent atteindre. Nous voyons que les utilités des deux agents sont croissantes avec la taille de la partition : les agents préfèrent donc avoir le plus d'information possible. De plus, la présence d'une asymptote montre également qu'à partir d'une certaine taille de partition, augmenter le nombre de cellules n'amènent plus à des gains significatifs. Ce qui illustre bien la différence avec la quantification classique, pour laquelle augmenter le nombre d'éléments dans la partition amènera toujours à une utilité plus importante.

30. Cet équilibre est appelé « Babbling equilibrium ».

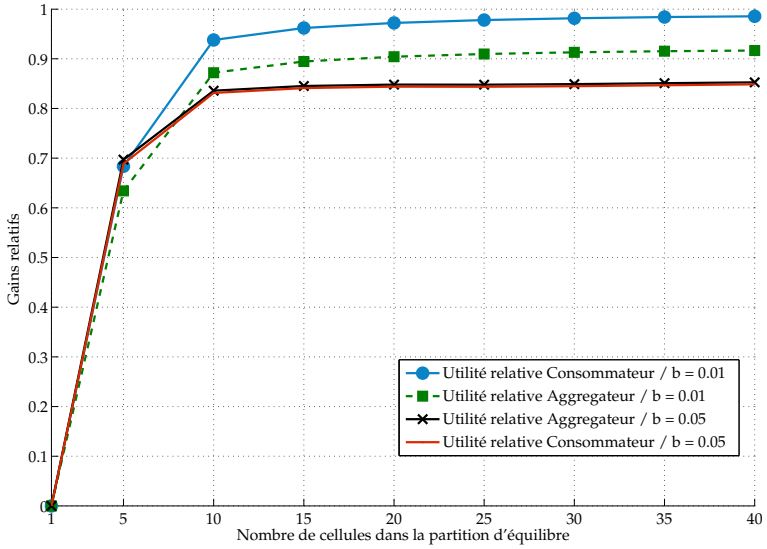

FIGURE 3.6 – **Gains relatifs des agents** : en fonction de la taille de la partition, nous traçons les utilités du consommateur et de l'agrégateur moyennées par l'utilité obtenue via une stratégie sans communication (ou de manière équivalente pour la partition d'équilibre à un élément). Ces utilités sont représentées pour deux valeurs du paramètre b : $b \in \{0.01, 0.05\}$.

3.4.2 Exemple de partitions et actions associées

Nous supposons ici que la capacité totale de charge du consommateur est de 22 kWh [31]. Une illustration d'une partition d'équilibre pour $M = 5$ messages est donnée en Figure 3.7. Nous avons choisi $b = 0.01$. Ce paramètre étant petit, les actions choisies par l'agrégateur se trouvent très proche du centre des cellules auxquelles elles appartiennent, ce qui se rapproche du modèle de quantification vectorielle.

Nous montrons maintenant l'influence de la valeur du paramètre b sur la valeur des actions choisies par l'agrégateur. Nous considérons en Figure 3.8 les valeurs suivantes : $b \in \{0.01, 0.05, 0.1\}$, et l'équilibre partitionnel de taille $M = 3$. Comme nous le voyons, plus b augmente, plus les actions de l'agrégateur ont tendance à se rapprocher de $(0, 0)$. En effet, l'augmentation du paramètre b traduit le fait que, pour l'agrégateur, le coût réseau (représenté par la fonction c) a de plus en plus d'importance par rapport au besoin du consommateur. Ce coût augmentant, l'agrégateur donnera de moins en moins de puissance de charge au consommateur. De plus, quelque soit b, pour cette configuration nous observons à chaque fois trois profils similaires : une action donnera une charge faible et égale sur les deux intervalles de temps (pour $b = 0.05$, le carré rouge correspondant à 4kWh en abscisse et en ordonnée), une action donnera une charge élevée sur le premier intervalle et une charge faible sur le deuxième (pour $b = 0.05$, le carré rouge correspondant à 11kWh en abscisse et 4.5kWh en ordonnée), et une dernière action donnera l'inverse

31. Cela correspond à la capacité d'une Renault ZOE : http://www.automobile-propre.com/voitures/renault-zoe/.

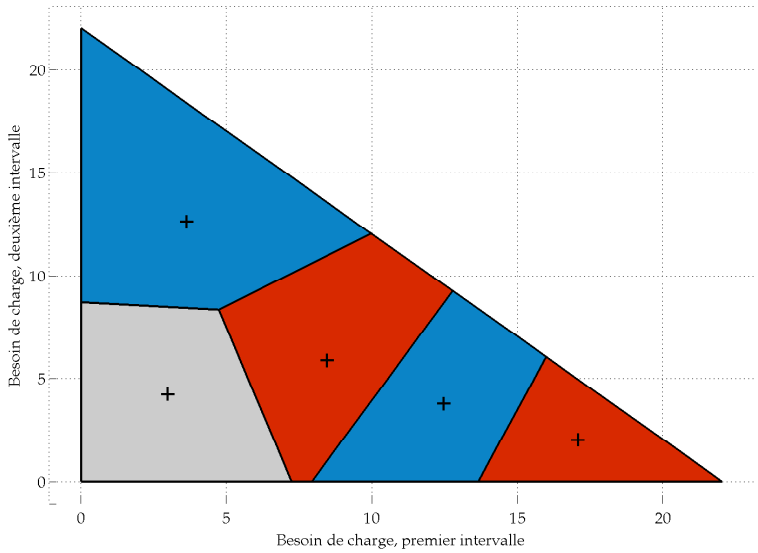

FIGURE 3.7 – **Exemple de partition d'équilibre en dimension 2** : nous représentons une partition d'équilibre pour $M = 5$ messages. Le simplexe est découpé en 5 cellules, les croix représentent les actions associées à chaque cellule de la partition d'équilibre, qui correspondent à deux charges effectives pour les deux intervalles de temps. Par exemple, si le consommateur envoie le message associé à la cellule grise, l'agrégateur lui donnera 4kWh sur le premier intervalle de temps et 3.5kWh sur le deuxième intervalle de temps. Pour cette illustration, le paramètre b vaut 0.01.

(pour $b = 0.05$, le carré rouge correspondant à 4.5kWh en abscisse et 11kWh en ordonnée). Si on augmente le nombre de messages disponibles, ce type de profils pourra toujours exister, mais d'autres apparaîtront de sorte que l'ensemble réalisable sera « quadrillé » plus finement.

3.5 Conclusion du chapitre

Dans ce chapitre, nous avons considéré un modèle simple de communication sans coût entre deux agents et avec une asymétrie d'information. Le point central de ce modèle est l'hypothèse d'utilités distinctes entre les deux agents, ce qui amène à étudier le problème d'un point de vue stratégique. Une des différences fondamentales avec la quantification est que toutes les ressources de communication ne sont pas utilisées pour des stratégies d'équilibre. Nous avons de plus une méthodologie pour obtenir de manière numérique les stratégies d'équilibre et les utilités associées pour des dimensions relativement grandes.
Ce travail peut être généralisé sur beaucoup d'aspects, comme par exemple d'autres fonctions d'utilité, d'autres croyances sur la distribution de probabilité de l'état, ou un nombre d'agents plus important. Néanmoins, le cadre étudié est général et permet d'établir de nouvelles connexions entre le modèle de jeu de signaux [27] et la quantification, et ouvre de nouveaux challenges qui concernent la généralisation de la quantification vectorielle à des utilités distinctes ainsi que le problème général de codage de source et de codage canal quand l'encodeur et le décodeur n'ont pas la même utilité. De plus, le modèle est réaliste et permettra d'obtenir des résultats numé-

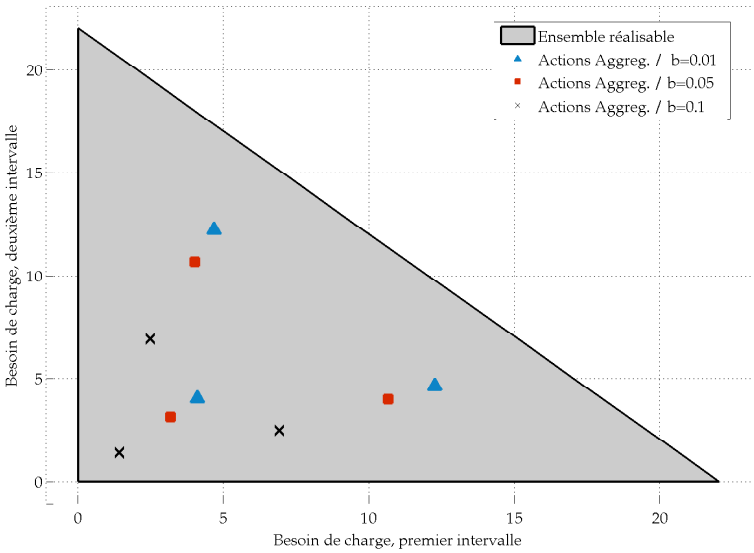

FIGURE 3.8 – **Charges effectives versus biais (dimension 2)** : le paramètre b appartient à $\{0.01, 0.05, 0.1\}$. Nous représentons les trois actions (de dimension 2) de l'agrégateur à l'équilibre en fonction du biais b, pour un besoin de charge sur deux intervalles de temps pour le consommateur.

riques intéressants pour de grandes dimensions, que nous avons illustré ici en dimension deux. Cette façon de communiquer permet de penser les futurs réseaux de distributions d'électricité de manière nouvelle et d'utiliser les ressources de communication de manière optimale. Par exemple, un scénario pratique typique modélise le besoin du consommateur sur un intervalle de temps de douze heures avec une prise de décision toutes les trente minutes.

Chapitre 4

Transformation d'une structure d'observation arbitraire à l'aide d'un encodeur

« I shall look at you out of the corner of my eye, and you will say nothing. Words are the source of misunderstandings. »
– Citation du Renard, *Le Petit Prince (1943)*, Antoine de Saint-Exupéry.

Dans ce chapitre, nous considérons un système avec un nombre d'agents quelconque supérieur à deux, qui jouent un jeu répété sur un grand nombre d'étapes. Au contraire des chapitres précédents, nous n'imposons pas de structure d'observation particulière et nous supposons que les agents observent les actions passées des autres agents via une structure d'information quelconque. Le problème que nous considérons ici est de savoir comment transformer cette structure d'observation quelconque en une autre structure d'observation qui permet de caractériser les utilités d'équilibre du jeu répété associé au système. Pour cela, nous nous demandons comment concevoir un encodeur qui permettrait à tous les agents d'obtenir une nouvelle structure d'observation. Nous nous concentrerons ici sur une structure qui garantit l'observation parfaite pour tous les agents des actions passées des autres agents. Cet encodeur, qui connaitra parfaitement les actions de tous les agents, aura pour but de compenser les erreurs que subissent les agents sur les observations des actions des autres agents. Nous cherchons à trouver le taux *minimal* qui garantit cela tout en étant robuste aux déviations d'un agent. Cela fera apparaître une nouvelle *contrainte d'information*.

4.1 Motivation et état de l'art

4.1.1 Introduction

Supposons que $K \geq 2$ agents jouent un jeu répété de long terme. La caractérisation des utilités d'équilibre d'un jeu répété est très liée aux hypothèses faites sur les capacités d'observations des agents. En effet, considérons par exemple la stratégie « grim trigger » : un agent coopère tant

qu'aucun agent ne trahit, mais dès lors qu'un agent dévie de la stratégie coopérative, l'agent utilisant « grim trigger » va sanctionner le déviateur en jouant une stratégie de punition sur toutes les étapes restantes du jeu. Cette stratégie est utilisée pour prouver le théorème Folk non parfait [75]. Elle n'est cependant possible que si une déviation d'un agent est détectée, ce qui suppose une certaine observation des actions passées des autres agents.
Dans ce chapitre, nous faisons l'hypothèse que les agents ont une structure d'observation quelconque, qui en toute généralité ne garantit pas de détecter les déviations. Nous ajoutons de plus un encodeur, qui permettra de transformer la structure d'observation initiale en une nouvelle structure que nous choisirons. Cet encodeur observe parfaitement les actions passées de tous les agents, que nous voyons comme une source d'information à transmettre. Cette source est plus générale que dans les chapitres précédents, car nous relâchons en partie l'hypothèse selon laquelle la distribution de probabilité de cette source est connaissance commune à tous les agents du système ainsi qu'à l'encodeur. Nous supposons que cette source d'information a K composantes et qu'une de ses composantes, la i-ème par exemple, a une distribution de probabilité connue seulement de l'Agent i. Cette hypothèse transcrit le fait qu'un agent peut dévier arbitrairement, et nous concevrons un code qui devra être robuste fasse à n'importe quelle distribution (déviation) utilisée par l'Agent i.
Nous nous focalisons sur les déviations unilatérales des agents puisque nous voulons reconstruire une structure d'observation qui permet d'avoir une caractérisation des utilités d'équilibre de Nash uniforme. Ce travail est un premier pas et une extension possible sera de considérer une coalition d'agents qui dévient de manière coopérative de sorte que la déviation bénéficie à tous les agents de la coalition. Le concept de solution qu'il faudra alors considérer est l'équilibre de Nash fort [32].

Il est important de noter que dans notre modèle l'encodeur ne cherche pas simplement à transmettre la source. Son rôle est en fait double : premièrement, le code conçu doit être fiable et doit permettre de garantir la nouvelle structure d'observation. Deuxièmement, ce code doit avoir une complexité minimale, c'est-à-dire que nous cherchons le taux optimal qui permet de passer de la structure d'observation quelconque à la nouvelle structure (après transformation). Les signaux reçus par les agents doivent être complémentaires de leurs signaux privés et la source d'information doit être envoyée de la manière la plus concise possible.
Le taux minimal sera caractérisé par une contrainte d'information, qui dépendra fortement de l'hypothèse faite sur la source concernant les possibilités de déviation des agents. Cette hypothèse amènera un coût en terme d'efficacité de compression. Elle dépendra également de la structure que nous voulons atteindre. Ici, nous nous concentrerons sur la manière d'assurer l'*observation parfaite* pour tous les agents. Retrouver l'observation parfaite pour tous les agents est très important d'un point de vue de la théorie des jeux, car grâce à cette hypothèse il est possible de caractériser les paiements d'équilibre d'un jeu répété de long terme. Cette condition suffisante permet d'obtenir un **théorème Folk** (voir le théorème B.34). Cela permet également d'avoir un ensemble d'équilibres contenant des équilibres efficaces, ce qui peut faire défaut lorsque les agents sont « aveugles » [3]. La condition d'observation parfaite est une condition forte. Une des extensions possibles serait de rétablir une condition moins contraignante, mais qui permet également d'obtenir un théorème Folk. Par exemple la condition « 2-connected observation graph » définie dans [74].

Dans ce chapitre, nous caractérisons la contrainte d'information qui permet de reconstruire l'observation parfaite virtuelle (Définition 4.3), en donnant une condition suffisante (Section

32. « Strong Nash Equilibrium ».

4.3.2) et nécessaire (Section 4.3.3). Cette condition permet de caractériser complètement la capacité du système considéré.
Nous appliquons enfin ce cadre à un exemple de jeu répété (Section 4.4).

4.1.2 Etat de l'art

Dans ce chapitre, nous généralisons le résultat de [60]. Dans cet article, les auteurs considèrent des jeux avec structure d'observation arbitraire. Les agents n'observent pas parfaitement les actions passées des autres agents, ils reçoivent simplement un signal privé via une structure d'observation fixée. L'ajout d'un encodeur qui aidera à rétablir l'observation parfaite est alors considéré, et cet encodeur enverra un signal *public*, qui sera observé de manière parfaite par tous les agents. L'encodeur est supposé observer parfaitement les actions de tous les agents ainsi que la structure d'observation, et a pour seul but d'améliorer les capacités d'observation des agents. Une condition suffisante sur le nombre de messages nécessaires pour reconstruire l'information est donnée. Cette condition est une contrainte d'information, et nous la généralisons [60] en supposant que l'encodeur choisit une certaine séquence qui sera l'entrée d'un canal de diffusion, et chaque agent reçoit une sortie de ce canal (bruité), qui sera son signal additionnel privé. Ceci est détaillé en Figure 4.1.
Dans ce modèle, nous considérons également qu'un des agents peut « dévier », c'est-à-dire que la distribution de probabilité sur sa séquence d'actions n'est connue que de lui. Ceci est différent par rapport au cas où la distribution de probabilité de la source est connaissance commune pour tous les agents du système. Ici, cette distribution peut varier arbitrairement. Les sources aux variations arbitraires sont étudiées en théorie de l'information, notamment dans [1, 2], nous permettant donc d'utiliser ce cadre comme cela été fait dans [60]. L'encodeur sera alors robuste au déviation d'un agent. Pour une source « classique », c'est-à-dire aux variations connues, une des principales techniques de preuve des théorèmes de codage est la méthode des types (voir Section A.3). Cette méthode suppose connue la distribution de probabilité de la source, elle n'est donc pas adaptée aux sources aux variations arbitraires, pour lesquelles d'autres techniques sont utilisées.
Nous supposons ici qu'un seul agent peut dévier. En effet, comme expliqué précédemment, la notion d'équilibre uniforme dans les jeux répétés s'intéresse uniquement aux déviations unilatérales, même si le jeu se déroule sur plusieurs étapes. Si nous voulons être robustes à des déviations de plusieurs agents, il faudra préciser la notion d'équilibre correspondante.

Le fait d'ajouter un canal de diffusion (voir Section A.5) amène des difficultés supplémentaires pour les schémas de codage. En effet, lorsque le nombre d'agents (d'utilisateurs) est $K = 2$, utiliser des codes source et canal séparés est optimal et il est donc possible de s'y restreindre pour atteindre la capacité du canal. Ce résultat est notamment utilisé dans [61] pour caractériser la capacité d'un système à deux utilisateurs et un espion. En revanche, pour un nombre d'utilisateur supérieur, $K > 2$, cela ne résout pas entièrement le problème, il y a une perte à séparer les deux codes. Or, pour un cas particulier du canal de diffusion avec une source commune pour tous les agents, il est prouvé dans [84] qu'utiliser un code source-canal *conjoint* amène à une caractérisation des taux (canal) atteignables. C'est un résultat important car c'est le premier pour lequel un code source-canal conjoint caractérise complètement les taux atteignables alors que la caractérisation pour des codages séparés est inconnue. Comme notre modèle correspond à ce cas particulier de canal de diffusion avec source commune, il n'est donc pas nécessaire de connaître la capacité du canal de diffusion général (avec messages privés). La capacité d'un tel

canal est encore inconnue à ce jour à part pour des cas particuliers comme par exemple le cas de dégradation physique [40].
Pour prouver que la contrainte d'information est suffisante, nous allons créer un code hybride qui utilisera le schéma de codage pour une source aux variations arbitraires (utilisé dans [60]) combiné au schéma de codage source-canal conjoint de Tuncel [84].

4.2 Description du modèle

Nous décrivons pour commencer la structure d'observation que nous étudions. Ensuite, nous faisons le lien entre ce modèle et le jeu répété qui lui est associé.

4.2.1 Structure d'observation avec assistance d'un encodeur

Nous considérons le modèle suivant, qui est représenté en Figure 4.1. Nous supposons un système à $K \geq 2$ agents et dénotons $\mathcal{A}_k$, $k \in \mathcal{K} \triangleq \{1, 2, \dots, K\}$ l'ensemble discret d'actions de l'Agent k, et $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \times \mathcal{A}_K$ l'ensemble discret des profils d'actions. Nous notons également $\mathcal{S}_k$ l'ensemble discret de signaux reçus par l'Agent k, et $\mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2 \times \cdots \times \mathcal{S}_K$. L'ensemble des entrées du canal sera noté $\mathcal{S}_0$, $|\mathcal{S}_0| < \infty$, et les ensembles des sorties du canal $\mathcal{V}_k$, $k \in \mathcal{K}$, $|\mathcal{V}_k| < \infty$. Nous notons enfin $\mathcal{V} = \mathcal{V}_1 \times \mathcal{V}_2 \times \cdots \times \mathcal{V}_K$.

Dans un souci de généralité, nous supposons que la longueur des séquences d'actions est n, et donc que la longueur des séquences de signaux privés reçus par les agents via la structure d'observation $\mathcal{T}$ est n également, mais la longueur des séquences d'entrées et sorties du canal de diffusion $\mathcal{T}_C$ sera m, avec $m \neq n$ en toute généralité. Nous posons alors :

$$\kappa \triangleq \frac{m}{n}. \tag{4.1}$$

Fixons maintenant la distribution de probabilité $Q_k^\star \in \Delta(\mathcal{A}_k)$ pour chaque agent $k \in \mathcal{K}$. Nous supposons donc qu'il n'y a pas de corrélations entre les agents ni de corrélations d'étape en étape pour un agent pris individuellement : la source que nous chercherons à transmettre est i.i.d.. Fixons également la structure d'observation (une probabilité conditionnelle) à

$$\mathcal{T} : \mathcal{A} \to \Delta(\mathcal{S}) \tag{4.2}$$

et une probabilité de transition pour le canal

$$\mathcal{T}_C : \mathcal{S}_0 \to \Delta(\mathcal{V}) \tag{4.3}$$

Lorsqu'une certaine séquence source fixée $a = (a_1, \dots, a_K) \in \mathcal{A}$ est tirée par rapport à la probabilité produit $Q^\star = Q_1^\star \times \cdots \times Q_K^\star \in \prod_{k=1}^K \Delta(\mathcal{A}_k)$, l'Agent k reçoit un symbole $s_k \in \mathcal{S}_k$ avec probabilité donnée par la probabilité conditionnelle :

$$\mathcal{T}(s_k|a) \triangleq \sum_{s_{-k} \in \mathcal{S}_{-k}} \mathcal{T}(s_k, s_{-k}|a). \tag{4.4}$$

De plus, pour une entrée du canal donnée s_0, l'Agent k reçoit un signal privé v_k via le canal de diffusion $\mathcal{T}_C$ selon la distribution de probabilité conditionnelle

$$\mathcal{T}_C(v_k|s_0) \triangleq \sum_{v_{-k} \in \mathcal{V}_{-k}} \mathcal{T}_C(v_k, v_{-k}|s_0) \tag{4.5}$$

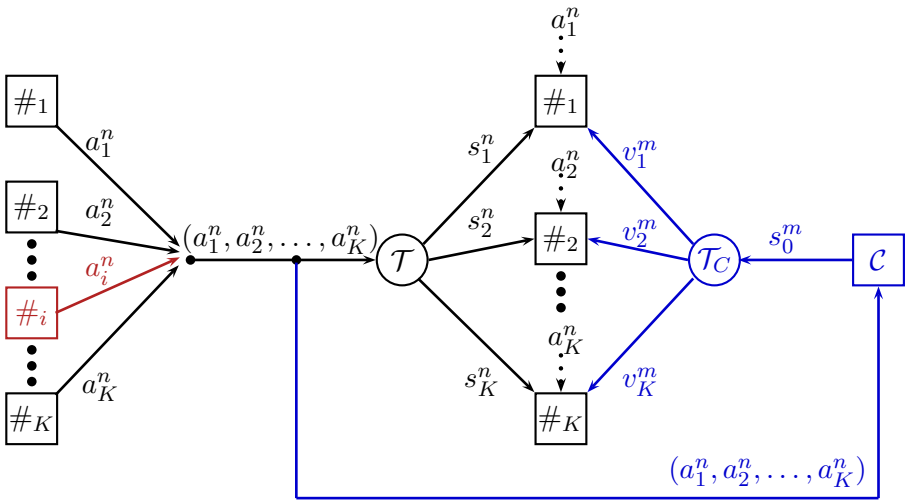

FIGURE 4.1 – **Structure d'observation avec encodeur** : chaque agent k (représenté deux fois sur ce schéma par $\#k$) choisit chaque élément de sa séquence d'actions a_k^n via la distribution de probabilité $Q_k^\star \in \Delta(\mathcal{A}_k)$ (nous supposons donc qu'il n'y a pas de corrélation entre les agents ni de corrélation d'étape en étape pour un agent pris individuellement), et doit reconstruire les séquences d'actions passées de tous les autres agents. Il observe un signal s_k^n reçu via la structure d'information arbitraire $\mathcal{T}$, et connaît sa propre séquence d'actions a_k^n (ceci est représenté via les flèches pointillées). De plus, un encodeur $\mathcal{C}$ (représenté en bleu) est ajouté au système pour rétablir l'observation parfaite pour tous les agents. L'encodeur aura pour rôle, en tenant compte des signaux déjà reçus par les agents, d'envoyer la source d'information de la manière la plus *concise* possible. Cet encodeur choisit la séquence d'entrée s_0^m du canal de diffusion $\mathcal{T}_C$ en fonction des séquences d'actions $(a_1^n, a_2^n, \ldots, a_K^n)$ qu'il observe parfaitement. Chaque agent k reçoit alors le signal additionnel v_k^m via le canal de diffusion $\mathcal{T}_C$, qui est supposé *orthogonal* au canal d'observation $\mathcal{T}$. Les signaux s_k^m et v_k^m sont donc indépendants. Nous supposons également qu'un certain agent i (représenté en couleur bordeaux) peut dévier et que donc ni les autres agents ni l'encodeur ne connaissent la distribution de probabilité utilisée par cet agent i pour choisir sa séquence d'actions. L'identité de cet agent déviateur est inconnue des (autres) agents ainsi que de l'encodeur, ce qui aura une influence sur le taux de compression optimal pour envoyer la source.

Le canal de diffusion est supposé *discret* et *sans mémoire*.
Nous supposons ici que les deux canaux $\mathcal{T}$ et $\mathcal{T}_C$ sont *orthogonaux*, et donc que les probabilités associées sont indépendantes. Une des extensions pourrait être de considérer un cas où il existe un unique canal avec deux entrées $a = (a_1, \ldots, a_K)$ et s_0, et K sorties (une pour chaque agent) $v_k, k = 1, 2, \ldots, K$.

Le scénario est le suivant : chaque agent k, $1 \leq k \leq K$, joue une certaine séquence d'actions i.i.d. a_k^n. Chaque profil d'actions du jeu $a^n = (a_1^n, a_2^n, \ldots, a_K^n)$ génère un profil de signaux $(s_1^n, s_2^n, \ldots s_K^n)$ via la distribution de probabilité conditionnelle $\mathcal{T}$. L'Agent k observe uniquement s_k^n de ce profil d'actions a^n. Nous ajoutons donc un encodeur $\mathcal{C}$, qui doit aider à la reconstruction pour tous les agents de a^n, en tenant compte d'une possible déviation d'un certain agent i, qui induit une incertitude sur la distribution de probabilité de la source. Nous faisons l'hypothèse ici

que l'identité du déviateur est *inconnue* de l'encodeur et des (autres) agents.
Cet encodeur, qui observe parfaitement le profil d'actions joués a^n, construit une séquence d'entrée s_0^m du canal de diffusion $\mathcal{T}_C$, de telle façon que chaque agent k reçoit une sortie du canal v_k^m. Le rôle de l'encodeur est que chaque agent reconstruise une *observation parfaite virtuelle* (Définition 4.3) à partir d'une séquence de paires de signaux (s_k^n, v_k^m) et la connaissance de sa propre séquence d'actions a_k^n.

4.2.2 Source aux variations arbitraires et *observation parfaite virtuelle*

Dans ce chapitre, nous supposons qu'un des agents peut changer arbitrairement sa stratégie de choix d'actions (sans que ces changements soient connaissance commune), donc que la distribution de probabilité de la source peut varier d'étape en étape (ou de profil d'actions en profil d'actions). Nous définissons précisément ceci d'un point de vue de la théorie de l'information grâce à la notion de **Source aux Variations Arbitraires** (SVA) :

Définition 4.1 (Source aux Variations Arbitraires (SVA))**.** *Soit $Q^\star \in \Delta(\mathcal{A})$ une distribution de probabilité (ici, une stratégie mixte pour un agent) et $\mathcal{V}_{ar}$ l'ensemble des états de la source :*

$$\mathcal{V}_{ar} = \cup_{k\in\mathcal{K}} \Delta(\mathcal{A}_k^\infty). \tag{4.6}$$

La source d'information aux variations arbitraires $a \in \mathcal{A}$ est dans un certain état $var \in \mathcal{V}_{ar}$, quand une des composantes du profil d'action a une distribution qui peut varier arbitrairement dans le temps et est complètement inconnue de l'encodeur.
Par exemple, pour toute longueur fixée n, quand une séquence d'états est $var = \mathcal{Q}_i \in \Delta(\mathcal{A}_i^\infty) \subset \mathcal{V}_{ar}$, la séquence d'actions $a^n = (a_1^n, \dots, a_K^n)$ est tirée selon la distribution de probabilité donnée par :

$$Q_{var}(a_1^n, \dots, a_i^n, \dots, a_K^n) = \left[Q_1^{\star\otimes n} \otimes \dots \otimes \mathcal{Q}_i \otimes \dots \otimes Q_K^{\star\otimes n} \right](a^n). \tag{4.7}$$

Remarque. *Les séquences produites par la source d'information sont de longueurs infinies, et non de longueur fini n. En effet, le codage est conçu de telle façon que la précision ϵ du codage (la valeur de la probabilité d'erreur maximale possible) dépend de la longueur des mots de code n, que nous ferons tendre vers l'infini. Or la source d'information ne doit pas dépendre de cette précision. En effet, si c'était le cas, il faudrait définir la source d'information pour chaque n différent. Nous voulons ici définir d'une seule et unique manière la source, à laquelle nous pourrons associer n'importe quelle précision ϵ.*

Dans ce contexte, nous considérons des stratégies compatibles avec l'SVA pour l'encodeur additionnel, définies comme suit :

Définition 4.2 (Code source-canal conjoint)**.** *Pour un canal de diffusion avec source commune, un code source-canal conjoint λ de tailles (n, m) pour l'encodeur $\mathcal{C}$ et les décodeurs $\mathcal{K}$ est constitué d'une fonction d'encodage $f_0^{(n,m)}$ et de K fonctions de décodage $(g_k)_{k\in\mathcal{K}}^{(n,m)}$ définies par :*

$$\begin{array}{rcccl} f_0^{(n,m)} & : & \mathcal{A}^n & \longrightarrow & \mathcal{S}_0^m \\ g_k^{(n,m)} & : & \mathcal{V}_k^m \times \mathcal{S}_k^n \times \mathcal{A}_k^n & \longrightarrow & \mathcal{A}^n, \qquad \forall k \in \mathcal{K}. \end{array} \tag{4.8}$$

Notons $\Lambda(n,m)$ l'ensemble des codes de longueurs fixées $n, m \in \mathbb{N}$.

$$\mathrm{P}_e^{(n,m)}(\lambda) = \sum_{k\in\mathcal{K}} \max_{i\in\mathcal{K}} \max_{\mathcal{Q}_i\in\Delta(\mathcal{A}_i^\infty)} \mathrm{P}_{\mathcal{Q}_i}(g_k^{(n,m)}(s_k^n, v_k^m, a_k^n) \neq a^n), \tag{4.9}$$

La probabilité d'erreur $\mathrm{P}_e^{(n,m)}(\lambda)$ du code code $\lambda \in \Lambda(n,m)$ est définie par l'équation (4.9) et correspond à la somme des probabilités d'erreurs de chaque décodeur $k \in \mathcal{K}$. Les maximums sur $\mathcal{Q}_i$ et i correspondent à la pire déviation possible $\mathcal{Q}_i \in \Delta(\mathcal{A}_i^\infty)$ du pire agent $i \in \mathcal{K}$ (respectivement), l'identité du déviateur i étant inconnue de l'encodeur et des décodeurs.

Notre but ici est de retrouver l'observation parfaite pour tous les agents, c'est-à-dire minimiser l'erreur sur le décodage des séquences d'actions passées pour tous les agents. Si cette erreur est minimisée, nous parlerons d'**Observation parfaite virtuelle**. En effet, cette observation est un peu différente car elle intervient avec un délai :

Définition 4.3 (Observation parfaite virtuelle (OPV))**.** *Fixons $\kappa \in \mathbb{R}$. Les agents $\mathcal{K}$ ont une observation parfaite virtuelle (OPV) de la source d'information $a \in \mathcal{A}$ si pour tout $\varepsilon > 0$, il existe un entier $\bar{n} \in \mathbb{N}$ tel que pour tout $n \geq \bar{n}$, il existe un code $\lambda \in \Lambda(n,m)$ avec $m = \kappa\, n$ tel que :*

$$\mathrm{P}_e^{(n,m)}(\lambda) \quad \leq \quad \varepsilon. \tag{4.10}$$

4.2.3 Lien avec les jeux répétés

Nous expliquons maintenant les liens entre la reconstruction de l'observation parfaite virtuelle et les stratégies de jeux répétés. Considérons le jeu répété avec signaux assisté d'un encodeur défini par

- le jeu

  $$(\mathcal{K}, (\mathcal{A}_k)_{k\in\mathcal{K}}, (u_k)_{k\in\mathcal{K}}) \tag{4.11}$$

  avec $u_k : \mathcal{A} \to \mathbb{R}$ la fonction d'utilité de l'Agent k, $k \in \mathcal{K}$;
- la structure d'observation $\mathcal{T}$ définie en équations (4.2) et (4.4) ;
- l'encodeur additionnel défini en équations (4.3) et (4.5).

L'encodeur est considéré comme un agent à part ici, et a pour seul but de reconstruire l'observation parfaite virtuelle.
Soit T l'horizon du jeu. Une stratégie de comportement pour l'Agent k, $k \in \mathcal{K}$ est une suite de fonction $(\sigma_k^t)_{t\in\{1,2,\dots,T\}}$ définie par :

$$\sigma_k^t : (\mathcal{A}_k \times \mathcal{S}_k \times \mathcal{V}_k)^{t-1} \to \Delta(\mathcal{A}_k) \tag{4.12}$$

Nous expliquons maintenant comment construire des stratégies d'équilibre qui tiennent compte du délai sur l'observation parfaite. Ce qui permet de retrouver les résultats des théorèmes folk.

Supposons que l'horizon T est découpé en B blocs de taille n : $T = nB$. Pour pouvoir communiquer, l'encodeur et les agents peuvent alors mettre en place un schéma de codage source-canal conjoint (Définition 4.2) avec $m = n$ qui garantit l'observation parfaite virtuelle (4.10). Ce schéma sera utilisé de bloc en bloc :

- L'encodeur va coder les séquences d'actions passées des agents du bloc b, $b \geq 2$, et l'envoyer sur le bloc $b+1$ via une fonction de la forme $f_0^{(n,n)}$ définie en équation (4.8). Il choisira une séquence arbitraire pour le premier bloc $b = 1$;
- A la fin du block $b+1$, chaque Agent k va décoder l'information du block b en utilisant une fonction de la forme $g_k^{(n,n)}$ définie en équation (4.8), en prenant en compte sa propre séquence d'actions passées, sa séquence de signaux $s_k^n(b+1)$ sur le bloc $b+1$ et sa séquence de signaux additionnels $v_k^n(b+1)$. Il y aura donc un délai de 2 blocs. Aucune stratégie de décodage n'est mise en place pour les deux premiers bloc $b = 1$ et $b = 2$. Chaque agent va ensuite faire un test statistique pour repérer une éventuelle déviation d'un agent. Ce test conditionnera le choix de la séquence d'actions sur le bloc $b+2$. Si le test statistique n'indique de déviation pour aucun des agents, les agents suivent sur toutes les étapes du bloc $b+2$ le plan principal de coopération défini au préalable du jeu. En revanche, si un déviateur est identifié, une stratégie de punition sera mise en place jusqu'à la fin du jeu. Ce schéma correspond à une stratégie de type « grim trigger ».

Nous expliquons dans la section suivante quelles contraintes doivent être satisfaite pour retrouver l'observation parfaite virtuelle.

4.3 Contraintes d'information pour retrouver l'observation parfaite

4.3.1 Graphe auxiliaire et coloriage

Dans ce cadre, nous devons transmettre de l'information de deux sortes : une où l'encodeur et les décodeurs connaissent la distribution de probabilité (pour les agents qui ne dévient pas), et l'autre où ils ne connaissent pas cette distribution (pour l'agent qui dévie et qui fait donc varier la source). Pour ce deuxième cas, nous avons besoin d'introduire des notions liées à la *théorie des graphes*. Nous commençons par définir un **graphe auxiliaire** :

Définition 4.4 (Graphe auxiliaire [60])**.** *Pour tout agent $i \in \mathcal{K}$, un graphe auxiliaire $\mathcal{G}_i$ est défini par $\mathcal{G}_i = (\mathcal{A}_i, \mathcal{E}_i)$:*

- *Les actions $a_i \in \mathcal{A}_i$ de l'agent $i \in \mathcal{K}$ sont les sommets du graphe.*
- *Il existe une arête $e_i = (a_i, a_i') \in \mathcal{E}_i$ entre deux actions $a_i \in \mathcal{A}_i$ et $a_i' \in \mathcal{A}_i$ si*

$$\exists\, a_{-i} \in Supp\, Q^\star_{-i},\ \exists\, k \in \mathcal{K},\ \exists\, s_k \in \mathcal{S}_k,\ \exists\, \delta > 0,\ \text{tels que :}$$
$$\min(\mathcal{T}(s_k|a_i, a_{-i}), \mathcal{T}(s_k|a_i', a_{-i})) \geq \delta \tag{4.13}$$

où $Supp\, Q^\star_{-i}$ est le support de la distribution de probabilité $Q^\star_{-i}$ définie par

$$Q^\star_{-i} = \bigotimes_{k \neq i} Q^\star_k \in \prod_{k \neq i} \Delta(\mathcal{A}_k). \tag{4.14}$$

Deux sommets $a_i \in \mathcal{A}_i$ et $a_i' \in \mathcal{A}_i$ sont donc voisins dans le graphe $\mathcal{G}_i$ si la probabilité que ces actions induisent, via $\mathcal{T}$, un même signal $s_k \in \mathcal{S}_k$ pour au moins un agent $k \in \mathcal{K}$ est strictement plus grande que 0.

Pour définir le **nombre chromatique** [16] du graphe $\mathcal{G}_i$, nous rappelons la définition de **coloriage** introduite dans [60] :

Définition 4.5 (Coloriage [60]). *Soit Φ_i un ensemble de couleurs. Un coloriage du graphe $\mathcal{G}_i$ est une fonction $\phi_i : \mathcal{A}_i \longrightarrow \Phi_i$ qui satisfait :*

$$\forall e_i = (a_i, a_i') \in \mathcal{E}_i : \ \phi_i(a_i) \neq \phi_i(a_i'). \tag{4.15}$$

Un coloriage minimal du graphe $\mathcal{G}_i$ est un coloriage ϕ_i pour lequel la cardinalité de l'ensemble des couleurs Φ_i est minimale. Le **nombre chromatique** χ_i du graphe $\mathcal{G}_i$ est la cardinalité $|\Phi_i|$ de l'ensemble des couleurs du coloriage minimal du graphe $\mathcal{G}_i$. C'est cette quantité qui sera utilisée dans les théorèmes de cette section.

Grâce à ces notions, nous pourrons coder l'information correspondant à la partie où la distribution de probabilité est inconnue de l'encodeur et des décodeurs. En effet, la quantité d'information utile pour l'encodeur $\mathcal{C}$ pour transmettre une composante séparément, qui correspond au déviateur i, et en tenant compte de l'information adjacente pour chaque décodeur $k \in \mathcal{K}$, est $\log_2 \chi_i$.
Pour l'information correspondant au cas où la distribution de probabilité est connue de l'encodeur et des décodeurs, nous utiliserons un schéma typique de théorie de l'information : le schéma de codage source-canal conjoint à travers un canal de diffusion [84]. Nous expliquons en détail les liens avec cet article dans la section suivante.

4.3.2 Condition suffisante pour l'*observation parfaite virtuelle*

Rappelons brièvement le théorème prouvé dans [84] :

Théorème 4.6 (Contrainte d'information pour le canal de diffusion [84]). *Pour une source discrète et sans mémoire A transmise à travers un canal de diffusion $P_{V_1 V_2 \dots V_K | S_0}$, avec information adjacente S_k pour tout agent $k \in \mathcal{K}$, communiquer de manière fiable est possible* ***en codage source-canal conjoint*** *si et seulement si :*

$$H(A|S_k) \leq \kappa I(S_0; V_k) \qquad \forall\, k \in \mathcal{K} \tag{4.16}$$

$$\Leftrightarrow \quad H(A|S_k) - \kappa I(S_0; V_k) \leq 0 \qquad \forall\, k \in \mathcal{K} \tag{4.17}$$

$$\Leftrightarrow \quad \max_{k \in \mathcal{K}} H(A|S_k) - \kappa I(S_0; V_k) \leq 0 \tag{4.18}$$

avec $\kappa = \frac{m}{n}$.

Voyons différentes configurations et leurs impacts sur la contrainte d'information.

- La capacité C_k du canal entre l'encodeur $\mathcal{C}$ et l'Agent k, $k \in \mathcal{K}$ est (Proposition A.14) :

$$C_k = \max_{\mathcal{P}_{S_0} \in \Delta(\mathcal{S}_0)} I(S_0; V_k) \tag{4.19}$$

- Commençons par le scénario le plus simple, en considérant le cas où aucun agent ne dévie (la source n'est donc plus soumise à des *variations arbitraires*), et aucun agent n'a d'information adjacente : $\mathcal{S}_k = \emptyset$ pour tout $k \in \mathcal{K}$. Dans ce cas, pour chaque agent k, le taux de la source d'information, c'est-à-dire de la variable aléatoire A privée de l'élément k (qui est déjà connu par l'agent k), doit être inférieur à l'information mutuelle entre la variable

aléatoire d'entrée S_0 du canal et la variable aléatoire V_k de sortie pour l'agent k [26]. La contrainte d'information associée sera alors :

$$\max_{k\in\mathcal{K}} H(A_{-k}) - \kappa\, I(S_0; V_k) < 0. \tag{4.20}$$

- Le deuxième cas généralise le premier en supposant qu'un certain Agent i, $i \in \mathcal{K}$ dévie et ses actions doivent être transmises séparément. Le codage symbole par symbole est optimal ici et la quantité d'information nécessaire pour encoder est $\log_2 |\mathcal{A}_i|$. Le terme $H(A_{-k})$ correspondant au taux source change et la contrainte d'information devient alors :

$$\max_{k\in\mathcal{K}} H(A_{-k,i}) + \log_2 |\mathcal{A}_i| - \kappa\, I(S_0; V_k) < 0. \tag{4.21}$$

- Si nous généralisons encore en ajoutant les informations adjacentes $(S_k(a_i), A_k)$ à chaque agent k, ce qui correspond au signal privé reçu par chaque agent via $\mathcal{T}$ et à sa propre séquence d'actions, le terme correspondant au taux va à nouveau changer. La quantité d'information nécessaire pour transmettre les actions de l'agent qui dévie i est maintenant $\log_2 \chi_i$, et la contrainte d'information devient :

$$\max_{k\in\mathcal{K}} H(A_{-k,i}|S_k(a_i), A_k) + \log_2 \chi_i - \kappa\, I(S_0; V_k) < 0. \tag{4.22}$$

Enfin, comme nous allons considérer la pire déviation possible pour le pire agent (la pire au sens de l'incertitude sur la variable aléatoire source), nous allons maximiser sur i et a_i :

$$\max_{\substack{i\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} \left[\max_{k\in\mathcal{K}} H(A_{-i,k}|S_k(a_i), A_k) + \log_2 \chi_i - \kappa\, I(S_0; V_k)\right] < 0 \tag{4.23}$$

En se basant sur ce raisonnement, nous allons prouver le résultat suivant :

Théorème 4.7 (Condition suffisante pour l'SVA avec signaux privés)**.** *Les agents $\mathcal{K}$ ont une observation parfaite virtuelle (OPV) pour la source aux variations arbitraires (SVA) $A \in \mathcal{A}$ si la condition suivante est vérifiée :*

$$\max_{\substack{i\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} \left[\max_{k\in\mathcal{K}} H(A_{-i,k}|S_k(a_i), A_k) + \log_2 \chi_i - \kappa\, I(S_0; V_k)\right] < 0, \tag{4.24}$$

où :

- *$A_{-i,k}$ est le profil d'actions sans les composantes i et k. Il est distribué selon $Q^\star_{-i,k} \in \prod_{\substack{j\neq i,\\ j\neq k}} \Delta(\mathcal{A}_j)$;*
- *$S_k(a_i)$ est le signal reçu par l'Agent k quand l'action a_i est fixée. Il est induit par A_{-i} et la probabilité de transition $\mathcal{T}$:*

$$\begin{aligned} \mathcal{T}_{a_i} : \mathcal{A}_{-i} &\longrightarrow \Delta(\mathcal{S}_k) \\ a_{-i} &\longrightarrow \mathcal{T}_{a_i}(s_k|a_{-i}) = \mathcal{T}(s_k|a_i, a_{-i}) \\ &= \sum_{s_{-k}\in\mathcal{S}_{-k}} \mathcal{T}(s_k, s_{-k}|a_i, a_{-i}). \end{aligned} \tag{4.25}$$

Remarque. *Les signaux S_k dépendent de l'action a_i ainsi que des autres actions a_{-i} à travers la probabilité conditionnelle $\mathcal{T}$. L'action a_i est supposée être la pire action que le déviateur i peut imposer au système de codage. Les autres actions sont tirées de manière i.i.d. avec la probabilité $Q^{\star}_{-i}$. Dans la démonstration, les actions du déviateur sont séparées des actions des autres agents. C'est en fixant chaque action du déviateur que nous repassons en « single letter ».*

Le premier maximum correspond au cas où le déviateur i choisit la pire action a_i en terme d'efficacité de codage pour le pire décodeur k (le deuxième maximum).
La démonstration de ce théorème est proposée en Annexe D.1. Nous résumons brièvement cette démonstration :

- La partie encodage est composée de trois étapes :

  1. Une première phase pour identifier l'Agent qui dévie. Pour cela, un test statistique sur les distributions de probabilité empirique est effectué, et le déviateur est celui qui a la distribution empirique la plus éloignée de sa distribution cible ;
  2. Une deuxième phase où l'encodeur doit coder (séparément) la séquence d'actions du déviateur i sans en connaître sa distribution. Pour cela, envoyer l'identité du déviateur ainsi que la séquence de couleurs associée à la séquence d'actions du déviateur est suffisant. Un codage source-canal séparé est effectué ;
  3. Une dernière étape où l'encodeur envoie les séquences d'actions correspondant à tous les agents sauf le déviateur. Il existe deux sous-cas ici, selon la fréquence d'utilisation d'un symbole. Si la fréquence n'est pas assez importante, l'utilisation d'une loi des grands nombres ne sera pas possible et les séquences correspondantes seront encodées de manière séparée, comme pour le déviateur i ; si cette fréquence est assez importante, nous utiliserons un codage de source-canal conjoint [84].

  La séquence s_0^m d'entrée du canal de diffusion est donc partitionnée en deux sous-séquences.

- Chaque décodeur k doit alors utiliser son information adjacente s_k^n et la séquence de signaux additionnels privés v_k^m pour décoder. Il devra également partitionner cette dernière et décoder premièrement l'identité du déviateur ainsi que sa séquence d'actions (via la séquence de couleurs), puis les séquences d'actions des autres agents.

- Les différentes probabilités d'erreurs de ce codage sont proches de 0 si la contrainte d'information (4.24) est vérifiée, ce que nous prouvons en utilisant des propriétés des ensembles typiques.

4.3.3 Condition nécessaire pour l'*observation parfaite virtuelle*

Pour la condition nécessaire, la difficulté ici est que nous devons gérer une partie avec une probabilité d'erreur proche de 0 (pour la partie où nous connaissons les distributions de probabilité source) et une partie avec probabilité d'erreur exactement égale à 0 (pour la partie source aux variations arbitraires, dont nous ne connaissons pas la distribution de probabilité). C'est sur cette partie où nous aurons besoin du graphe auxiliaire et du nombre chromatique associé à ce graphe.

Théorème 4.8 (Condition nécessaire pour l'SVA avec signaux privés)**.** *Supposons que les agents $\mathcal{K}$ ont une observation parfaite virtuelle (OPV) pour la source aux variations arbitraires (SVA) $A \in \mathcal{A}$. Alors la condition suivante est vérifiée :*

$$\max_{\substack{i\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} \max_{k\in\mathcal{K}} \left[H(A_{-i,k}|A_k, S_k(a_i)) + \log_2 \chi_i - \kappa I(S_0; V_k)\right] < 0 \tag{4.26}$$

La démonstration de ce théorème se trouve en Annexe D, Section D.2.

4.4 Application aux jeux répétés

Considérons le dilemme du prisonnier répété avec matrice de paiements donné dans le Tableau 4.2. Nous avons donc $|\mathcal{A}| = 4$, et nous fixons $|\mathcal{S}_0| = |\mathcal{V}_k| = 3$, $k = 1, 2$. Notre but est de voir quelles sont les stratégies mixtes compatibles avec la contrainte d'information (4.18). Pour simplifier, nous supposons $\kappa = 1$ (ou de manière équivalente $n = m$).

	L	R
T	(3, 3)	(0, 4)
B	(4, 0)	(1, 1)

Tableau 4.2 – Matrice de paiements du Dilemme du prisonnier.

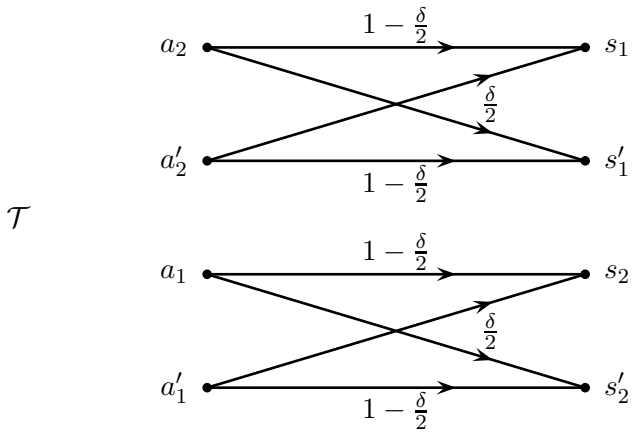

FIGURE 4.3 – **Structure d'observation privée $\mathcal{T}$** qui dépend du paramètre $\delta \in [0, 1]$. Deux signaux privés s_1 et s_1' sont observés par l'Agent 1 avec une certaine probabilité conditionnelle par rapport aux deux actions de l'Agent 2 a_2 et a_2'. Il en est de même pour s_2 et s_2' pour l'Agent 2.

Dans l'esprit de [60], nous considérons la même structure d'observation $\mathcal{T}$ décrite en Figure 4.3 : avec probabilité $1-\frac{\delta}{2}$, $\delta \in [0; 1]$, l'Agent k, $k = 1, 2$, observe le bon signal s_k et avec probabilité $\frac{\delta}{2}$ il observe le mauvais signal s_k'. Dans cette configuration, les deux structures d'observation (celles de chaque agent) sont orthogonales.

Pour la partie canal de diffusion, nous considérons la structure d'observation décrite en Figure 4.4 : avec probabilité $1-\frac{\alpha}{2}$, $\alpha \in [0;1]$, l'Agent k, $k=1,2$, reçoit la bonne sortie du canal, avec probabilité $\frac{\alpha}{2}$ il reçoit la mauvaise sortie du canal, et ne reçoit pas la troisième sortie (probabilité 0). Dans cette configuration, les deux structures d'observation sont également orthogonales.

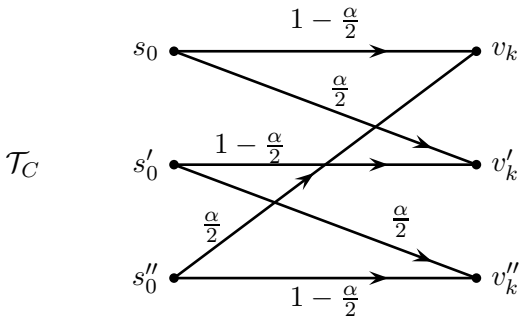

FIGURE 4.4 – **Structure du canal de diffusion $\mathcal{T}_C$** qui dépend du paramètre $\alpha \in [0,1]$.

Dans ce contexte, la contrainte d'information (4.24) peut être réécrite de manière à être utilisée pour des simulations. Cette formulation est explicitée en Section D.3 de l'Annexe D.

Nous considérons ici trois valeurs pour le paramètre δ : $\delta \in \{0.2, 0.31, 0.35, 1\}$. Pour le paramètre du canal de diffusion, nous prenons $\alpha = 0.01$. Nous comparons les régions d'utilité obtenues avec la contrainte d'information (4.24) et celle obtenue dans [60], qui est :

$$\max_{\substack{i\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} \left[\max_{k\in\mathcal{K}} H(A_{-i,k}|S_k(a_i), A_k) + \log_2 \chi_i \right] - \log_2 |S_0| < 0 \tag{4.27}$$

La comparaison est faite en Figure 4.5.

Nous voyons que les régions d'utilité avec la contrainte (4.24) sont plus petites que celles avec la contrainte (4.27), qui est une condition moins restrictive. Le fait d'avoir un bruit sur le canal de diffusion réduit la taille des régions d'utilité.

4.5 Conclusion du chapitre

Nous avons étudié ici un jeu muni d'une structure d'information arbitraire, où chaque agent reçoit un signal privé via une structure d'observation fixée ainsi qu'un signal additionnel privé via l'ajout d'un encodeur qui utilise un canal de diffusion. L'encodeur observe lui parfaitement les actions passées de tous les agents. L'encodeur aide à la reconstruction de l'observation parfaite pour tous les agents. Nous avons étendu les résultats de [60].
D'autres extensions peuvent être considérées, comme par exemple rétablir la condition « 2-connected observation graph », qui est moins restrictive que la condition d'observation parfaite. Considérer des déviations de plusieurs agents et non d'un unique agent, et étudier l'ensemble des équilibres forts du jeu répété est une extension à étudier avec le formalisme mis en place dans ce chapitre. Nous pouvons également penser à enlever l'hypothèse d'orthogonalité entre le canal d'observation $\mathcal{T}$ et le canal de diffusion additionnel $\mathcal{T}_C$, et supposer par exemple que chaque agent ne reçoit qu'un seul et unique signal privé qui dépend à la fois du canal d'observation fixé et des choix de l'encodeur. Enfin, une des extensions possibles serait de considérer des stratégies de comportement pour les agents. C'est-à-dire que les éléments des séquences d'actions a_k^n, $k \in \mathcal{K}$, seront dépendants les uns des autres.

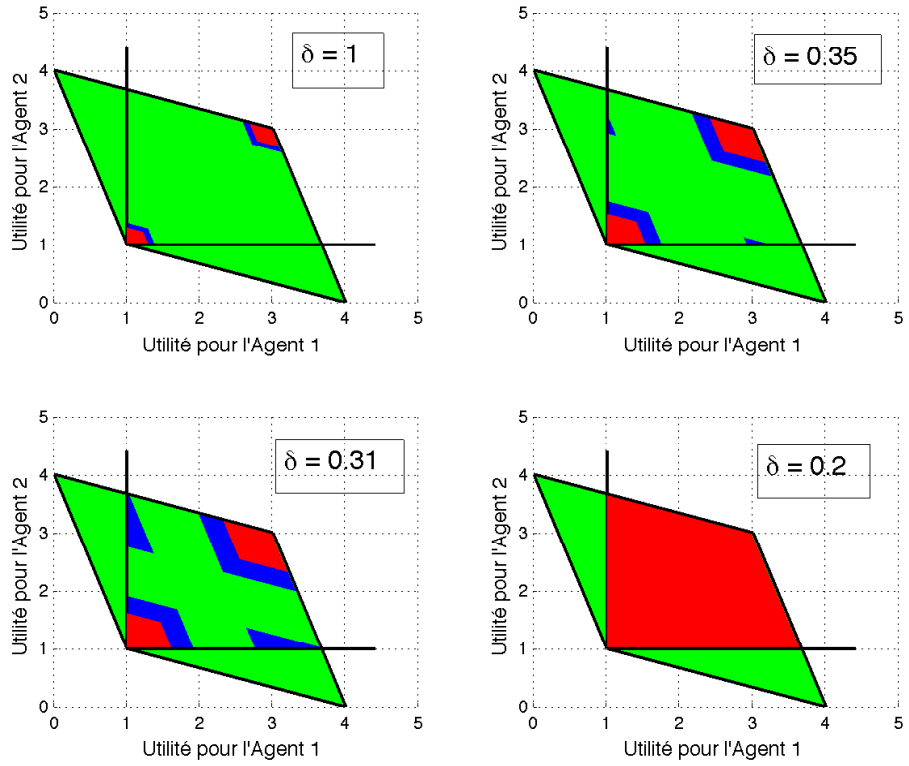

FIGURE 4.5 – **Comparaison des régions d'utilités** entre la contrainte d'information (4.24) (en rouge) et la contrainte (4.27) (en bleu) pour les valeurs des paramètres $\delta \in \{0.2, 0.31, 0.35, 1\}$ et $\alpha = 0.01$. L'enveloppe convexe des utilités possibles est en verte. Les traits noirs représentent les niveaux minmax.

Chapitre 5

Conclusion et perspectives

« Nobody ever figures out what life is all about, and it doesn't matter. Explore the world. Nearly everything is really interesting if you go into it deeply enough. »

– Richard P. Feynman.

Le but principal de cette thèse était l'étude de différentes structures d'information et des moyens d'optimiser l'utilisation des ressources d'information en fonction de ces structures. Pour cela, plusieurs modèles ont été étudiés : un modèle à deux agents, intérêts communs et communication coûteuse ; un modèle à deux agents, intérêts divergents et communication sans coût ; ainsi qu'un modèle avec un nombre d'agents quelconques où l'essentiel pour les agents est d'avoir l'observation parfaite des actions passées des autres agents. Dans ces modèles, nous avons supposé une asymétrie d'information : un agent a de l'information sur l'état du système qu'un autre agent n'a pas. Nous avons donc caractérisé, pour ces modèles, la manière optimale de coordonner les différents agents du système. Pour cela, nous avons utilisé des outils de la théorie de l'information et en particulier nous avons fait du codage de source. En plus du challenge théorique de ces problèmes, nous avons appliqué nos résultats à des cas pratiques de communication sans fil, notamment au problème de contrôle de puissance. Notre nouvelle approche, le *contrôle de puissance codé*, est complètement générale et nous montrons qu'elle permet d'obtenir le bon compromis entre transmission d'information et maximisation de l'utilité. Elle apparaît comme très prometteuse par rapport à des algorithmes classiques utilisés jusqu'alors.
Nous avons également développé une application aux réseaux de véhicules électriques. Notre approche sur un modèle de base permet de bien comprendre les enjeux et amène des pistes de réflexion pour une gestion intelligente de l'information dans ces réseaux.

Ces modèles sont relativement simples mais sont essentiels à la compréhension des mécanismes d'optimisation des ressources de communication, et permettent d'avoir une vision de base pour étudier des modèles plus généraux. En effet, en plus d'amener une nouvelle approche de résolution de problème d'optimisation et de théorie des jeux via des outils de théorie de l'information, cela permet de penser d'une manière différente les algorithmes et codes qui seront implémentés en pratique pour des problèmes de communication sans fil.

Les perspectives de cette thèse sont d'étudier des modèles plus généraux pour pouvoir se rapprocher au mieux de modèles pratiques réalistes et renforcer un peu plus les liens entre deux des théories scientifiques les plus importantes de ces dernières décennies. Nous pouvons par exemple considérer les extensions suivantes :

- Augmenter le nombre d'agents pour les modèles des chapitres 2 et 3 paraît essentiel. En augmentant ce nombre, il faudra redéfinir l'asymétrie d'information et préciser la connaissance qu'ont les agents sur l'état du réseau ;
- Pour les modèles avec deux agents puis dans un deuxième temps pour des modèles à $K \in \mathbb{N}$ agents, nous pouvons supposer que chaque agent possède une information partielle sur l'état du système et étudier une structure d'information bidirectionnelle. Le modèle de théorie de l'information de Shannon [80] ainsi que les modèles de théorie des jeux [73],[77],[7] ou [9] pourront servir de base de travail ;
- Nous pouvons également considérer un modèle avec un espion et caractériser la nouvelle contrainte qui permettra d'être robuste à l'espionnage ;
- Ajouter plusieurs phases de communication plutôt que l'envoi d'un unique message pourra également être une extension à considérer. Ceci permettra de réduire au moins partiellement le biais entre les utilités des agents ;
- Enfin, passer d'espaces discrets à des espaces continus est un axe de travail qui est à l'étude et qui donnera lieu à une soumission à la conférence de contrôle européenne (ECC 2015). Le passage du discret au continu doit être fait avec attention, à commencer par la définition des entropies différentielles. Cet article permet de faire le lien entre discret et continu. Nous appliquons de plus ces résultats au problème bien connu de Witsenhausen.

D'un point de vue pratique, en plus des applications en communication sans fil et aux Smart Grids correspondant aux extensions théoriques ci-dessus, nous pensons qu'il est possible d'appliquer ces modèles à d'autres contextes notamment pour la transmission d'information en économie ou au sein d'une entreprise pour optimiser la gestion de l'information. En conclusion, nous pouvons affirmer que ce domaine qui regroupe des problèmes à la croisée de deux théories n'en est qu'à ses débuts et est très prometteur.

Annexe A

Théorie de l'information

« You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one really knows what entropy really is, so in a debate you will always have the advantage. »

– Suggestion de John von Neumann à Claude Shannon à propos du nom à donner à sa nouvelle « mesure d'incertitude », *Scientific American 224, (September 1971).*

A.1 Définitions et propositions de base

Avant d'aller dans les notions propres à la théorie de l'information, nous faisons un détour par la théorie des probabilités et définissons les notions suivantes :

Définition A.1 (Distance totale de variation)**.** *La distance de variation totale entre deux distributions de probabilité P et Q définies sur un ensemble $\mathcal{X}$ est la moitié de la distance L_1 entre ces deux distributions, définie par :*

$$||P - Q||_{VT} \triangleq \sum_{x \in \mathcal{X}} \frac{1}{2}|P(x) - Q(x)| \tag{A.1}$$

Définition A.2 (Distance de Hamming)**.** *Soit $\mathcal{A}$ un alphabet et F l'ensemble des suites de longueur n à valeur dans $\mathcal{A}$. La distance de Hamming d_H entre deux éléments $\underline{a} = (a_i)_{1 \leq i \leq n}$ et $\underline{b} = (b_i)_{1 \leq i \leq n}$ de F est le nombre d'éléments de l'ensemble des images de $\underline{a}$ qui diffèrent de celle de $\underline{b}$:*

$$d_H(a, b) = \#\{i : a_i \neq b_i\} \tag{A.2}$$

Nous commençons donc maintenant la théorie de l'information à proprement parlé et définissons les notions d'*Entropie*, d'*Entropie Conditionnelle*, ainsi que certaines propositions de bases sur ces notions :

Définition A.3 (Entropie). *L'entropie d'une variable aléatoire Z de distribution de probabilité Q sur l'ensemble fini $\mathcal{Z}$ est notée $H_Q(Z)$ ou $H(Z)$ s'il n'y a pas d'ambiguïté sur la distribution de probabilité considérée, et est définie par :*

$$H_Q(Z) = H(Z) \triangleq -\sum_{z\in\mathcal{Z}} Q(z)\log_2 Q(z). \tag{A.3}$$

Définition A.4 (Entropie conditionnelle). *L'entropie conditionnelle d'une variable aléatoire Z par rapport à une variable aléatoire Y lorsque ces deux variables ont pour distribution conjointe de probabilité Q sur $(\mathcal{Y}\times\mathcal{Z})$ est notée $H_{Q_{YZ}}(Z|Y)$ ou $H(Z|Y)$ s'il n'y a pas d'ambiguïté sur la distribution de probabilité considérée, et est définie par :*

$$H_Q(Z|Y) = H(Z|Y) \triangleq -\sum_{(y,z)\in(\mathcal{Y}\times\mathcal{Z})} Q(y,z)\log_2 \frac{Q(y,z)}{\sum_{z'\in\mathcal{Z}} Q(y,z')}. \tag{A.4}$$

L'entropie mesure l'incertitude d'une variable aléatoire, et l'entropie conditionnelle mesure l'incertitude d'une variable aléatoire en tenant compte de la connaissance d'une autre variable aléatoire (possiblement corrélée à la première). Cette mesure est bornée comme le montre la proposition suivante :

Proposition A.5 (Borne sur l'entropie). *Pour toute variable aléatoire Z prenant ses valeurs dans l'ensemble fini $\mathcal{Z}$:*

$$0 \leq H(Z) \leq \log_2 |\mathcal{Z}| \tag{A.5}$$

Une des propriétés de base qui lie l'entropie et l'entropie conditionnelle est la suivante :

Proposition A.6 (Règle de chaîne pour l'entropie). *Pour une distribution conjointe Q sur $(\mathcal{Y}\times\mathcal{Z})$ et la distribution marginale $Q_Y \triangleq \sum_{z\in\mathcal{Z}} Q(y,z)$, l'équation suivante est toujours vérifiée :*

$$H_Q(Y,Z) = H_{Q_Y}(Y) + H_Q(Z|Y) \tag{A.6}$$

L'entropie et l'entropie conditionnelle sont aussi liées par la notion d'information mutuelle :

Définition A.7 (Information mutuelle). *L'information mutuelle entre deux variables aléatoires Y et Z lorsque ces deux variables ont pour distribution conjointe de probabilité Q sur $(\mathcal{Y}\times\mathcal{Z})$ est notée $I_Q(Y;Z)$ ou $I(Y;Z)$ s'il n'y a pas d'ambiguïté sur la distribution de probabilité considérée, et est définie par :*

$$I_Q(Y;Z) \triangleq \sum_{(y,z)\in(\mathcal{Y}\times\mathcal{Z})} Q(y,z)\cdot\log_2 \frac{Q(y,z)}{\sum_{z'\in\mathcal{Z}} Q(y,z')\cdot\sum_{y'\in\mathcal{Y}} Q(y',z)}, \tag{A.7}$$

L'information mutuelle est une mesure de corrélation moyenne entre deux variables aléatoires ayant une certaine distribution conjointe de probabilité. Les notions d'entropie et d'information mutuelle sont connectées via la proposition suivante :

Proposition A.8 (Information mutuelle et entropie)**.**

$$\begin{aligned} I_Q(Y;Z) &= H_{Q_Y}(Y) - H_Q(Y|Z) && \text{(A.8)} \\ &= H_{Q_Z}(Z) - H_Q(Z|Y) && \text{(A.9)} \\ &= H_{Q_Y}(Y) + H_{Q_Z}(Z) - H_Q(Y,Z) && \text{(A.10)} \end{aligned}$$

L'information mutuelle a la propriété d'être toujours positive (pour des variables aléatoires *discrètes*), ce qui nous amène à la proposition suivante :

Proposition A.9 (Le conditionnement réduit l'entropie)**.**

$$I_Q(Y;Z) \geq 0 \Leftrightarrow H_Q(Y|Z) \leq H_{Q_Y}(Y) \tag{A.11}$$

Nous avons aussi la notion d'information mutuelle conditionnelle :

Définition A.10 (Information mutuelle conditionnelle)**.** *L'information mutuelle conditionnelle entre deux variables aléatoires Y et Z conditionnellement à une troisième variable X lorsque ces trois variables ont pour distribution conjointe de probabilité Q sur $(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$ est notée $I_Q(Y;Z|X)$ ou $I(Y;Z|X)$ s'il n'y a pas d'ambiguïté sur la distribution de probabilité considérée, et est définie par :*

$$I_Q(Y;Z|X) \triangleq \sum_{(x,y,z)\in(\mathcal{X}\times\mathcal{Y}\times\mathcal{Z})} Q(x,y,z) \cdot \log_2 \frac{Q(x,y,z)\cdot Q_X(x)}{\sum_{z'\in\mathcal{Z}} Q(x,y,z') \cdot \sum_{y'\in\mathcal{Y}} Q(x,y',z)} \tag{A.12}$$

avec

$$Q_X(x) \triangleq \sum_{(y,z)\in(\mathcal{Y}\times\mathcal{Z})} Q(x,y,z) \tag{A.13}$$

Ainsi qu'une règle de chaîne pour l'information mutuelle :

Proposition A.11 (Règle de chaîne pour l'information mutuelle)**.**

$$I_Q(Y;Z,X) = I_{Q_{YZ}}(Y;Z) + I_Q(Y;Z|X) \tag{A.14}$$

Une dernière proposition utile dans les démonstrations réciproques notamment, est la suivante :

Proposition A.12 (Inégalité de traitement des données ou « Data-processing inequality »)**.** *Soit X, Y et Z trois variables aléatoires qui vérifient la chaîne de Markov*

$$X \to Y \to Z \tag{A.15}$$

Alors l'inégalité suivante est vérifiée :

$$I(X;Y) \geq I(X;Z). \tag{A.16}$$

A.2 Théorèmes de codage

Nous commençons par les propositions de base, à savoir le taux de compression source minimal ainsi que la capacité maximale d'un canal.

Un encodeur $\mathcal{C}$ observe une séquence de réalisations $s^n \in \mathcal{S}^n$ d'une source d'information et l'envoie au décodeur $\mathcal{D}$ en utilisant un ensemble de messages $\mathcal{M}$. Le décodeur reconstruit une séquence $\hat{s}^n \in \mathcal{S}^n$ à partir du message reçu. Le débit d'information moyen d'un code est alors mesuré par la quantité

$$R \triangleq \frac{\log_2 M}{n} \tag{A.17}$$

où M est la cardinalité de l'ensemble $\mathcal{M}$.
Un débit est atteignable s'il est plus grand que R et que le code qui induit ce débit rend la probabilité d'erreur de l'évènement $\hat{s}^n \neq s^n$ arbitrairement faible. La distribution de probabilité de la source est supposée connue par l'encodeur et le décodeur. Shannon a prouvé le résultat suivant :

Proposition A.13 (Taux source [79])**.** *Soit S une variable aléatoire source de distribution de probabilité $Q_S \in \Delta(\mathcal{S})$ sur l'ensemble fini $\mathcal{S}$. Alors le débit R d'un code associé à cette source d'information et qui rend la probabilité d'erreur arbitrairement faible, vérifie :*

$$R \geq H_{Q_S}(S). \tag{A.18}$$

La capacité canal est en un sens le problème inverse d'un taux source. L'encodeur utilise cette fois une séquence d'entrées du canal pour transmettre un certain nombre de messages $M = |\mathcal{M}|$. Le décodeur observera alors la séquence de sorties du canal et reconstruira le message envoyé par l'encodeur. Le débit d'information moyen d'un code est, comme pour un débit d'un code source, $R = \frac{\log_2 M}{n}$. L'objectif d'un code canal est de maximiser ce débit tout en assurant une probabilité d'erreur arbitrairement faible.

Proposition A.14 (Capacité d'un canal [79])**.** *Soit un canal $\mathcal{T} : \mathcal{X} \to \Delta(\mathcal{Y})$ pour des ensembles finis $\mathcal{X}$ et $\mathcal{Y}$. Supposons qu'il soit utilisé de manière indépendante d'étape en étape (sans mémoire). Alors le débit R d'un code associé à cette source d'information et qui rend la probabilité d'erreur arbitrairement faible, vérifie :*

$$R \leq C \triangleq \max_{Q_X \in \Delta(\mathcal{X})} I_{Q_{XY}}(X;Y) \tag{A.19}$$

avec $Q_{XY} = Q_X \cdot \mathcal{T}$.
C est la capacité du canal.

Ces deux propositions se démontrent en créant un code via les séquences typiques.

Nous pouvons alors étudier un système où une source d'information peut être transmise à travers un canal. Nous parlerons alors de code source-canal conjoint. Un des résultats importants en théorie de l'information est que concaténer un code source et un code canal indépendants l'un de l'autre est optimal pour transmettre une source discrète à travers un canal discret et sans mémoire :

Théorème A.15 (Théorème de séparation source-canal)**.** *Une source discrète S de distribution de probabilité $Q_S \in \Delta(\mathcal{S})$ peut être transmise à travers un canal $\mathcal{T} : \mathcal{X} \to \Delta(\mathcal{Y})$ discret et sans mémoire si et seulement si*

$$H_{Q_S}(S) \leq \max_{Q_X \in \Delta(\mathcal{X})} I_{Q_{XY}}(X;Y). \tag{A.20}$$

Ce résultat de séparation n'est pas vrai dans toutes les configurations de source et de canal possible. Nous prouvons dans le Chapitre 2 que séparer le codage source et le codage canal pour le système considéré n'induit pas de perte en terme de performance. En revanche, dans le modèle du Chapitre 4, nous utilisons un code source-canal conjoint car nous savons que la séparation ferait perdre le code en efficacité.

A.3 Séquences typiques

Une des techniques les plus utiles dans cette thèse est la *méthode des types* [28]. Nous définissons ici ce qu'est le type d'une séquence.

Définition A.16 (Type [26])**.** *Soit $T \geq 1$. Pour toute séquence de réalisations z^T de la variable aléatoire générique Z, le type de z^T, dénoté T_{z^T}, est la distribution de probabilité sur $\mathcal{Z}$ définie par*

$$\mathrm{T}_{z^T}(z) \triangleq \frac{N(z|z^T)}{T} = \frac{1}{T}\sum_{t=1}^{T} \mathbb{1}_{\{z_t = z\}}, \tag{A.21}$$

où la notation $N(z|z^T)$ signifie le nombre de positions de z^T ayant la lettre z.

Viens ensuite l'ensemble des séquences typiques, c'est-à-dire des suites qui ont une distribution empirique proche d'une distribution donnée.

Définition A.17 (Ensemble de séquences typiques [40])**.** *Soit Q_Z une distribution de probabilité sur $\mathcal{Z}$. L'ensemble des séquences ϵ-typiques de longueur n par rapport à Q_Z est*

$$\mathcal{T}_\epsilon^n(Q_Z) \triangleq \{z^n \in \mathcal{Z}^n : \forall z \in \mathcal{Z},\ |T_{z^n}(z) - Q_Z(z)| < \epsilon\}. \tag{A.22}$$

Cette définition peut être étendue pour un couple de suites.

Définition A.18 (Ensemble de suites conjointement typiques [40])**.** *Soit Q_{YZ} une distribution de probabilité sur $(\mathcal{Y} \times \mathcal{Z})$. L'ensemble des séquences conjointement ϵ-typiques de longueur n par rapport à Q_{YZ} est*

$$\mathcal{T}_\epsilon^n(Q_{YZ}) \triangleq \{(y^n, z^n) \in (\mathcal{Y} \times \mathcal{Z})^n : \forall (y,z) \in (\mathcal{Y} \times \mathcal{Z}),\ |T_{y^n,z^n}(y,z) - Q_{YZ}(y,z)| < \epsilon\}. \tag{A.23}$$

Avec cette définition, nous pouvons alors définir l'ensemble des suites conditionnellement ϵ-typiques :

Définition A.19 (Ensemble des suites conditionnellement typiques [40])**.** *L'ensemble des suites conditionnellement ϵ-typiques par rapport à la séquence $y^n \in \mathcal{Y}^n$ est défini par :*

$$\mathcal{T}_\epsilon^n(Q_{YZ}|y^n) \triangleq \{z^n \in \mathcal{Z}^n : (y^n, z^n) \in \mathcal{T}_\epsilon^n(Q_{YZ})\} \tag{A.24}$$

A.4 Lemmes pour les théorèmes de codage

Dans cette section, nous donnons les définitions précises de lemmes tirés de [40]. Ces lemmes sont d'une grande importance dans les théorèmes d'atteignabilité en général et en particulier dans ceux des chapitres 2 et 4. Ils donnent notamment des conditions sur la cardinalité des livres de codes pour que la probabilité d'erreur du code soit arbitrairement proche de 0.

Lemme A.20 (Lemme de recouvrement)**.** *Soit $(U, X, \hat{X}) \sim Q_{UX\hat{X}}$, avec $Q_{UX\hat{X}}$ une distribution conjointe fixée. Soit $\epsilon > \epsilon' > 0$. Soit (U^n, X^n) une paire de séquences aléatoires telles que*

$$\lim_{n\to\infty} Pr\{(U^n, X^n) \in \mathcal{T}_{\epsilon'}^n(Q_{UX})\} = 1 \tag{A.25}$$

Définissons les séquences aléatoires $\hat{X}^n(w)$, $w \in \mathcal{A}$ où $|\mathcal{A}| \geq 2^{nR}$ conditionnellement indépendantes les unes des autres et de X^n sachant U^n, toutes distribuées selon $\prod_{i=1}^n Q_{\hat{X}|U}(\hat{x}_i|u_i)$. Alors il existe $\delta(\epsilon)$ qui tend vers 0 avec $\epsilon \to 0$ tel que

$$\lim_{n\to\infty} Pr\{(U^n, X^n, \hat{X}^n(w)) \notin \mathcal{T}_\epsilon^n(Q_{UX\hat{X}}) \text{ pour tout } w \in \mathcal{A}\} = 0 \tag{A.26}$$

si

$$R > I_{Q_{UX\hat{X}}}(X; \hat{X}|U) + \delta(\epsilon). \tag{A.27}$$

Lemme A.21 (Lemme de typicité conditionnelle)**.** *Définissons les variables aléatoires $(X, Y) \sim Q_{XY}$. Supposons que $x^n \in \mathcal{T}_{\epsilon'}^n(Q_X)$ et $Y^n \sim p(y^n|x^n) = \prod_{i=1}^n Q_{Y|X}(y_i|x_i)$. Alors, pour tout $\epsilon > \epsilon' > 0$,*

$$\lim_{n\to\infty} Pr\{(x^n, Y^n) \in \mathcal{T}_\epsilon^n(Q_{XY}\} = 1. \tag{A.28}$$

Lemme A.22 (Packing Lemma)**.** *Soit $(U, X, Y) \sim Q_{UXY}$ trois variables aléatoires. Soit $(U^n, Y^n) \sim p(u^n, y^n)$ une paire de séquences générées selon une distribution arbitraire, et pas forcément selon $\prod_{i=1}^n Q_{U,Y}(u_i, y_i) = \sum_x \prod_{i=1}^n Q_{UXY}(u_i, x, y_i)$. Soit $X^n(w)$, $w \in \mathcal{A}$, où $|\mathcal{A}| \leq 2^{nR}$, des séquences aléatoires, chacune distribuée selon $\prod_{i=1}^n Q_{X|U}(x_i|u_i)$. Supposons enfin que pour tout $w \in \mathcal{A}$:*

$$Pr\{X^n(w), Y^n|U^n\} = Pr\{X^n(w)|U^n\} \cdot Pr\{Y^n|U^n\}. \tag{A.29}$$

Une séquence $X^n(w)$ peut être arbitrairement dépendante des autres séquences $X^n(w')$, $w' \neq w$.

Alors, il existe $\delta(\epsilon)$ qui tend vers 0 avec $\epsilon \to 0$ telle que

$$\lim_{n\to\infty} P\left\{\bigcup_{w\in\mathcal{A}} (U^n, X^n(w), Y^n) \in \mathcal{T}_\epsilon^n(Q_{UXY})\right\} = 0 \tag{A.30}$$

si

$$R < I_{Q_{UXY}}(X; Y|U) - \delta(\epsilon). \tag{A.31}$$

Lemme A.23 (Lemme de typicité conjointe). *Définissons les trois variables aléatoires* $(X, Y, Z) \sim Q_{XYZ}$ *avec* Q_{XYZ} *une distribution de probabilité fixée. Définissons également* $\epsilon > \epsilon' > 0$.

1. *Supposons que* $(x^n, y^n) \in (\mathcal{X} \times \mathcal{Y})^n$ *est une paire de séquences arbitraires (tirées selon n'importe quelle distribution conjointe sur* $(\mathcal{X} \times \mathcal{Y})^n$*), et supposons que* Z^n *vérifie*

$$Pr\{Z^n|x^n, y^n\} = Pr\{Z^n|x^n\} \tag{A.32}$$

et est i.i.d. :

$$Z^n \sim \prod_{i=1}^{n} Q_{Z|X}(\cdot|x_i) \tag{A.33}$$

Alors il existe $\delta(\epsilon) > 0$ *telle que* $\delta(\epsilon) \to_{\epsilon \to 0} 0$ *et :*

$$Pr\left\{(x^n, y^n, Z^n) \in \mathcal{T}_\epsilon^n(Q_{XYZ})\right\} \leq 2^{-n(I(Y;Z|X)-\delta(\epsilon))}. \tag{A.34}$$

2. *Supposons que* $(x^n, y^n) \in \mathcal{T}_{\epsilon'}^n(Q_{XY})$ *et que* Z^n *vérifie*

$$Pr\{Z^n|x^n, y^n\} = Pr\{Z^n|x^n\} \tag{A.35}$$

et est i.i.d. :

$$Z^n \sim \prod_{i=1}^{n} Q_{Z|X}(\cdot|x_i) \tag{A.36}$$

Alors il existe $\delta(\epsilon) > 0$ *telle que* $\delta(\epsilon) \to_{\epsilon \to 0} 0$ *et, pour* n *suffisamment grand :*

$$Pr\left\{(x^n, y^n, Z^n) \in \mathcal{T}_\epsilon^n(Q_{XYZ})\right\} \geq 2^{-n(I(Y;Z|X)+\delta(\epsilon))}. \tag{A.37}$$

Lemme A.24 (Lemme de Markov). *Supposons que la chaîne de Markov* $X - Y - Z$ *est vérifiée par trois variables aléatoires* X*,* Y *et* Z*. Si, pour un couple de séquences donné* $(x^n, y^n) \in \mathcal{T}_\epsilon^n(Q_{XY})$*,* Z^n *est tirée selon* $\prod_{i=1}^n Q_{Z|Y}(z_i|y_i)$*, alors*

$$Pr\{(x^n, y^n, Z^n) \in \mathcal{T}_\epsilon^n(Q_{XYZ})\} \geq 1 - \delta(\epsilon). \tag{A.38}$$

avec $\delta(\epsilon) \to 0$ *as* $n \to \infty$.

Enfin, nous finissons par un des théorèmes les plus utilisés dans les théorèmes de codage pour la partie nécessaire. Nous en avons besoin en particulier dans le Chapitre 4. Ce théorème permet de borner la probabilité que l'estimation $\hat{X}$ de la variable source X soit différente de celle ci.

Théorème A.25 (Inégalité de Fano). *Pour les variables aléatoires* $X, \hat{X} \in \mathcal{X}$ *et* $Y \in \mathcal{Y}$ *qui vérifient la chaîne de Markov* $X - Y - \hat{X}$*, et en posant* $P_e \triangleq Pr(X \neq \hat{X})$*, nous avons :*

$$H(P_e) + P_e \log|\mathcal{X}| \geq H(X|\hat{X}) \geq H(X|Y). \tag{A.39}$$

Cette inégalité implique :

$$1 + P_e \log|\mathcal{X}| \geq H(X|Y). \tag{A.40}$$

A.5 Deux modèles de canaux en théorie de l'information

Nous proposons dans cette section deux modèles de canal qui sont très importants dans cette thèse. Nous commençons par le canal de diffusion :

Définition A.26 (Canal de diffusion)**.** *Un canal de diffusion est un canal de communication avec un encodeur $\mathcal{C}$ et K décodeurs $\mathcal{D}_k$, $K \geq 2$, $k \in \{1, 2, \ldots, K\}$. La source est composée d'un message commun M_0 et d'un message privé M_k pour chaque décodeur $k \in \{1, 2, \ldots, K\}$. L'encodeur choisit une séquence X^n d'entrées du canal de diffusion $\mathcal{T}(y_1, y_2, \ldots, y_K|x)$, chaque décodeur k reçoit une séquence de sortie Y_k^n, et doit décoder le message commun M_0 et son message privé M_k. Ceci est illustré en Figure A.1.*

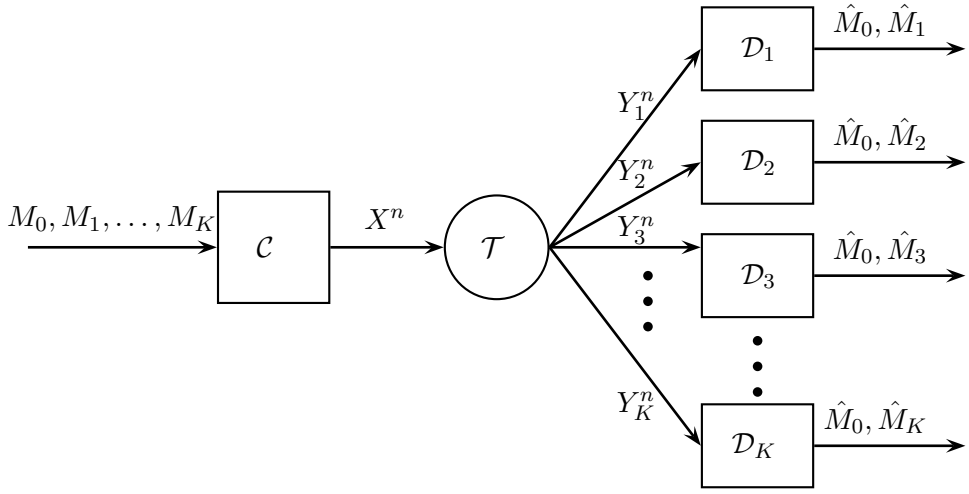

FIGURE A.1 – **Canal de diffusion**.

Nous avons également le canal à accès multiple, qui est représenté sur la figure suivante :

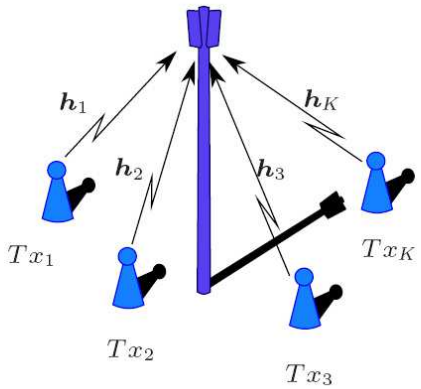

FIGURE A.2 – **Canal à accès multiple** : plusieurs agents Tx_k, $k \in \{1, 2, \ldots, K\}$ se connectent au même canal de communication. Chaque agent k, $k \in \{1, 2, \ldots, K\}$ choisit une entrée X_k^n du canal $\mathcal{T}(y|x_1, x_2, \ldots, x_K)$. La sortie du canal est Y^n.

Annexe B

Théorie des jeux

« Bernoulli's real contribution was to coin a word. The word has been translated into English as utility. It describes this subjective value people place on money. »
– William Poundstone, *Fortune's Formula : The Untold Story of the Scientific Betting System That Beat the Casinos And Wall Street (2006).*

Dans cette annexe, nous présentons quelques définitions et propositions utiles pour cette thèse. Pour une étude détaillée de la théorie des jeux, le lecteur peut se référer au livre [64].

B.1 Quelques définitions utiles

Nous commençons par définir ce qu'est un jeu en un coup ainsi qu'un jeu répété. Ces définitions sont cruciales notamment pour le Chapitre 3.

Définition B.27 (Jeu sous forme normale)**.** *Un jeu sous forme normale est un vecteur ordonné*

$$\mathcal{G} \triangleq (\mathcal{K}, (\mathcal{S}_k)_{k\in\mathcal{K}}, (u_k)_{k\in\mathcal{K}}) \tag{B.1}$$

avec

- $\mathcal{K} = (1, 2, \dots, K)$ *l'ensemble fini des joueurs ;*
- $\mathcal{S} = (\mathcal{S}_1 \times \mathcal{S}_2 \times \cdots \times \mathcal{S}_K)$ *le produit cartésien des ensembles finis de stratégies* S_k *de chaque joueur* $k \in \mathcal{K}$ *;*
- $(u_k)_{k\in\mathcal{K}} = (u_1, u_2, \dots, u_K)$ *le vecteur de fonctions d'utilité* u_k *de chaque joueur* $k \in \mathcal{K}$ *avec*

$$u_k : \mathcal{S} \longrightarrow \mathbb{R} \tag{B.2}$$

Chaque joueur $k \in \mathcal{K}$ *choisit une action* s_k *dans l'ensemble* $\mathcal{S}_k$ *et reçoit l'utilité* $u_k(s_1, s_2, \dots, s_K)$ *qui dépend des choix d'actions de tous les joueurs.*

Les joueurs peuvent choisir d'utiliser telle ou telle action avec une certaine probabilité, plutôt que de n'utiliser qu'une action fixe. Cela nous mène à la notion de *stratégie mixte* :

Définition B.28 (Stratégies mixtes)**.** *Une stratégie mixte pour le joueur $k \in \mathcal{K}$ est une distribution de probabilité P_k sur son ensemble de stratégie pure $\mathcal{S}_k$. Si tous les joueurs utilisent des stratégies mixtes, les fonctions d'utilités sont définies par*

$$u_k : \Delta(\mathcal{S}) \longrightarrow \mathbb{R} \qquad \forall\, k \in \mathcal{K} \tag{B.3}$$

et nous considérons l'espérance sur $P = (P_1 \times P_2 \times \cdots \times P_K)$.

Lorsque le nombre de joueurs est 2, et le nombre de stratégies pour chaque joueur est suffisamment petit, un jeu sous forme normale peut être représenté par une matrice comportant S_1 lignes et S_2 colonnes, avec S_1 le nombre de stratégies du joueur 1 et S_2 le nombre de stratégies du joueur 2. Nous parlerons alors de **jeu sous forme matricielle**. Par exemple, nous avons :

Définition B.29 (Exemple de jeu sous forme matricielle)**.** *Le jeu Faucon et Colombe peut être représenté sous la forme matricielle suivante*

$J_1 \diagdown J_2$	Colombe	Faucon
Colombe	(2, 2)	(10, 0)
Faucon	(0, 10)	(-5, -5)

Tableau B.1 – Matrice de paiements pour le jeu Faucon et Colombe.

où les deux joueurs ont deux actions possibles (Colombe, Faucon), le joueur 1 choisissant une ligne et le joueur 2 une colonne. Les utilités du joueur 1 sont donnés par la première coordonnée de chaque couple, ceux du joueur 2 par la deuxième.

Dans ces jeux, les joueurs font un seul choix de stratégie et obtiennent un paiement. Une extension est de considérer plusieurs choix successifs pour chaque joueur, nous parlons alors de **jeu répété**. Nous définissons tout d'abord les stratégies associées aux jeux répétés T fois. Pour cela, nous avons besoin de la notion d'*histoire d'un jeu* :

Définition B.30 (Histoire d'un jeu)**.** *L'information disponible à l'étape $t+1$ d'un jeu est les actions jouées par tous les joueurs lors des t premières étapes du jeu. Pour tout $t \geq 0$, l'ensemble des* histoires en t-étapes *est défini par*

$$H(t) \triangleq \mathcal{S}^t = \underbrace{\mathcal{S} \times \mathcal{S} \times \cdots \times \mathcal{S}}_{t \text{ times}} \tag{B.4}$$

Pour $t = 0$, nous identifions $H(0) \triangleq \{\emptyset\}$, qui ne contient aucune action.

Une stratégie de comportement pour le joueur k est un plan d'actions qui associe une stratégie mixte à jouer après toute histoire du jeu possible :

Définition B.31 (Stratégie de comportement)**.** *Une* stratégie de comportement *pour un joueur k dans un jeu répété T fois est une fonction qui associe une stratégie mixte avec toute histoire de longueur plus petite que T*

$$f_k : \bigcup_{t=0}^{T-1} H(t) \to \Delta(\mathcal{S}_k) \tag{B.5}$$

L'ensemble des stratégies de comportement du joueur k d'un jeu répété T fois est dénoté $\mathcal{B}_k^T$.

Nous pouvons maintenant définir un jeu répété :

Définition B.32 (Jeu répété)**.** *Soit $\mathcal{G} = (\mathcal{K}, (\mathcal{S}_k)_{k\in\mathcal{K}}, (u_k)_{k\in\mathcal{K}})$ un jeu de base. Le jeu répété T fois, est le jeu*

$$\mathcal{G}_T \triangleq (\mathcal{K}, (\mathcal{B}_k^T)_{k\in\mathcal{K}}, (\gamma_k^T)_{k\in\mathcal{K}}) \tag{B.6}$$

avec $\gamma_k^T(s) \triangleq \frac{1}{T}\sum_{t=1}^T u_k(s)$ l'utilité moyenne du joueur k sur T étapes.

Ces jeux sont à information complète *car les joueurs observent parfaitement les choix d'actions passées des autres joueurs. Une généralisation à l'information incomplète peut être faite en supposant que les joueurs reçoivent uniquement des signaux sur les actions passées des autres joueurs.*

B.2 Equilibre de Nash et théorème Folk

Nous donnons maintenant la définition de l'Equilibre de Nash, un des concepts de solutions les plus importants en théorie des jeux. L'équilibre de Nash est une notion de stabilité : pour des stratégies d'équilibre, aucun joueur n'a intérêt à changer de stratégie pendant que les autres joueurs jouent la stratégie d'équilibre.

Définition B.33 (Équilibre de Nash)**.**

- *Un vecteur de stratégies $s^* = (s_1^*, s_2^*, \dots, s_K^*)$ est un* équilibre de Nash *d'un jeu $\mathcal{G} = (\mathcal{K}, (\mathcal{S}_k)_{k\in\mathcal{K}}, (u_k)_{k\in\mathcal{K}})$ si pour tout joueur $k \in \mathcal{K}$ et pour toute stratégie $s_k \in \mathcal{S}_k$, l'équation suivante est satisfaite :*

$$u_k(s^*) \geq u_k(s_k, s_{-k}). \tag{B.7}$$

  Le vecteur d'utilité $u(s^*)$ est le vecteur d'utilité d'équilibre associé à l'équilibre de Nash s^*.

- *Pour un jeu répété T fois $\mathcal{G}_T$, un vecteur de stratégies $\tau^* = (\tau_1^*, \tau_2^*, \dots, \tau_K^*)$ est un équilibre de Nash, si pour tout joueur $k \in \mathcal{K}$ et pour toute stratégie $\tau_k \in \mathcal{B}_k^T$, l'équation suivante est satisfaite :*

$$\gamma_k^T(\tau^*) \geq \gamma_k^T(\tau_k, \tau_{-k}). \tag{B.8}$$

  Le vecteur d'utilité $\gamma^T(s^*)$ est le vecteur d'utilités d'équilibre du jeu $\mathcal{G}_T$.

Une autre notion importante est la notion de **niveau minmax** : pour un certain joueur k, c'est la pire valeur qu'il peut obtenir. Tous les joueurs excepté le joueur k choisissent des stratégies qui minimisent l'utilité maximum du joueur k. Le *niveau minmax* est défini par :

$$\overline{v}_k \triangleq \min_{s_{-k}\in\otimes_{i\neq k}\Delta(\mathcal{S}_i)} \max_{s_k\in\Delta(\mathcal{S}_k)} u_k(s_k, s_{-k}). \tag{B.9}$$

Définissons enfin l'ensemble des vecteurs d'utilité V par

$$V \triangleq \{x \in \mathbb{R}^K : x_k \geq \overline{v}_k \text{ pour tout joueur } k \in \mathcal{K}\}. \tag{B.10}$$

Cet ensemble est l'ensemble des vecteurs d'utilité pour lesquels chaque joueur reçoit au moins sa valeur minmax. Il est appelé l'ensemble des *paiements individuellement rationnels.*

Nous présentons maintenant un **Théorème Folk** pour un jeu répété un nombre fini de fois, qui expose que sous certaines conditions techniques, l'ensemble des paiements d'équilibre est (ou approche) l'ensemble des paiements réalisables et individuellement rationnels du jeu de base.

Théorème B.34 (Un théorème Folk)**.** *Supposons que pour tout joueur $k \in \mathcal{K}$, il existe un équilibre $s^*(k)$ du jeu de base $\mathcal{G} = (\mathcal{K}, (\mathcal{S}_k)_{k\in\mathcal{K}}, (u_k)_{k\in\mathcal{K}})$ tel que*

$$u_k(s^*(k)) > \overline{v}_k. \tag{B.11}$$

Alors, pour tout $\epsilon > 0$, il existe $T_0 \in \mathbb{N}$ tel que pour tout $T \geq T_0$ et pour tout vecteur de paiements réalisables et individuellement rationnels $x \in F \cap V$, il existe un équilibre τ^* du jeu répété T fois $\mathcal{G}_T$ qui implique des paiements ϵ-proche de x (en norme max) :

$$||\gamma^T(\tau^*) - x||_\infty < \epsilon. \tag{B.12}$$

avec $F \triangleq$ conv $\{u(s), s \in \mathcal{S}\}$ l'enveloppe convexe des vecteurs d'utilité d'étape, qui contient les vecteurs d'utilités moyennes.

Il existe plusieurs théorèmes Folk, en fonction du type de jeu considéré. Une des hypothèses importantes pour prouver ce théorème est l'hypothèse d'observation parfaite des actions passées de tous les joueurs.

B.3 Dynamique de meilleure réponse

En théorie des jeux, la **meilleure réponse** est une stratégie qui produit le meilleur résultat pour un joueur, en prenant pour fixées les stratégies des autres joueurs. Pour l'équilibre de Nash, chaque joueur joue une meilleure réponse aux stratégies des autres joueurs.
En *théorie évolutive des jeux*, une branche de la théorie des jeux, il existe la notion de **dynamique de meilleure réponse**, qui est un processus itératif où à chaque itération, les joueurs mettent à jour leur stratégie et jouent leur meilleure réponse aux stratégies des autres joueurs de l'étape précédente. Il existe des conditions grâce auxquelles ce processus converge éventuellement vers un équilibre de Nash du jeu.
Nous utilisons cette dynamique au Chapitre 3.

Annexe C

Démonstrations du chapitre 3

« Do not worry about your difficulties in Mathematics. I can assure you mine are still greater. »

– Albert Einstein, *answer to a little girl who had difficulties in school with mathematics (letter, 1943).*

C.1 Preuve de la proposition 3.1

Démonstration. Considérons un nombre de messages M fixé et les fonctions d'utilités de la forme (3.9) et (3.10) pour $T = 1$, c'est-à-dire :

$$U_C((s_m)_m, (a_m)_m) = -\sum_{m=1}^{M} \int_{s_{m-1}}^{s_m} (a_m - s)^2 \, p(s) \, ds \tag{C.1}$$

$$U_A((s_m)_m, (a_m)_m) = -\sum_{m=1}^{M} \int_{s_{m-1}}^{s_m} \left[(a_m - s)^2 + bc(a_m)\right] \, p(s) \, ds \tag{C.2}$$

Pour prouver l'existence d'un équilibre de Nash mixte, il suffit de prouver la continuité de ces fonctions [33] par rapport à $((s_m)_{m\in\{0,1,\dots,M\}}, (a_m)_{m\in\{1,2,\dots,M\}})$. Nous la prouvons pour la fonction $U_A(\cdot,\cdot)$, ce qui impliquera la continuité de $U_C(\cdot,\cdot)$ en prenant $b = 0$.

Soit une suite $((s^i_m)_{m\in\{0,1,\dots,M\}}, (a^i_m)_{m\in\{1,2,\dots,M\}})_i$ telle que $s^i_m \to s_m$ quand $i \to \infty$ pour tout $m \in \{0, \dots, M\}$ et $a^i_m \to a_m$ quand $i \to \infty$ pour tout $m \in \{1, \dots, M\}$.
Alors, nous savons qu'à partir d'un certain rang, s^i_m est proche de s_m, et en particulier plus proche de s_m que ne le sont s_{m-1} et s_{m+1}. Dès lors, nous avons deux cas : soit l'état est dans l'intervalle m de la partition $(s_m)_m$ et de la partition $(s^i_m)_m$, soit l'état est dans l'intervalle m de la partition s_m mais dans l'intervalle $m-1$ ou $m+1$ de la partition s^i_m. C'est ce que traduit l'équation (C.4) , en faisant attention au cas extrême en 0 et en 1.

Supposons de plus que pour tout m : $s_m^i \leq s_m$. Cette hypothèse ne fait pas perdre en généralité et les autres cas pourront être traité de la même manière. Nous avons :

$$\begin{aligned}
&\Big|U_A((s_m)_m,(a_m)_m) - U_A((s_m^i)_m,(a_m^i)_m)\Big| \\
&= \Big|\sum_{m=1}^{M}\int_{s_{m-1}}^{s_m}\left[(a_m-s)^2+bc(a_m)\right]\ p(s)\,ds - \int_{s_{m-1}^i}^{s_m^i}\left[(a_m^i-s)^2+bc(a_m^i)\right]\ p(s)\,ds\Big| \qquad \text{(C.3)} \\
&= \underbrace{\Big|\int_0^{s_1^i}\left[(a_1-s)^2+bc(a_1)-(a_1^i-s)^2-bc(a_1^i)\right]\ p(s)\,ds\Big|}_{(1)} \\
&\quad + \underbrace{\Big|\sum_{m=1}^{M-1}\int_{s_m^i}^{s_m}\left[(a_{m-1}-s)^2+bc(a_{m-1})-(a_m^i-s)^2-bc(a_m^i)\right]\ p(s)\,ds\Big|}_{(2)} \\
&\quad + \underbrace{\Big|\sum_{m=2}^{M-1}\int_{s_{m-1}}^{s_m^i}\left[(a_m-s)^2+bc(a_m)-(a_m^i-s)^2-bc(a_m^i)\right]\ p(s)\,ds\Big|}_{(3)} \\
&\quad + \underbrace{\Big|\int_{s_{M-1}}^{1}\left[(a_M-s)^2+bc(a_M)-(a_M^i-s)^2-bc(a_M^i)\right]\ p(s)\,ds\Big|}_{(4)} \qquad \text{(C.4)}
\end{aligned}$$

Par définition, $\exists\, i_0 \in \mathbb{N}$ tel que $\forall i \geq i_0$: $|s_m - s_m^i| < \epsilon_s$ pour tout $m \in \{0,\dots,M\}$ et $|a_m - a_m^i| < \epsilon_a$ pour tout $m \in \{1,\dots,M\}$. Nous avons alors, pour tout $i \geq i_0$:

$$(1) \leq \delta_1(\epsilon_a) \qquad \text{(C.5)}$$

avec $\delta_1(\epsilon_a) \to 0$ quand $\epsilon_a \to 0$, par continuité de la fonction $a \mapsto (a-s)^2 + bc(a)$.
De même :

$$(3) \leq \delta_3(\epsilon_a) \qquad \text{(C.6)}$$

avec $\delta_3(\epsilon_a) \to 0$ quand $\epsilon_a \to 0$ et

$$(4) \leq \delta_4(\epsilon_a) \qquad \text{(C.7)}$$

avec $\delta_4(\epsilon_a) \to 0$ quand $\epsilon_a \to 0$.

Pour (2), nous intégrons une fonction continue et bornée, donc, comme la borne inférieure d'intégration tend vers la borne supérieure d'intégration :

$$(2) \leq \delta_2(\epsilon_s) \qquad \text{(C.8)}$$

avec $\delta_2(\epsilon_s) \to 0$ quand $\epsilon_s \to 0$.

Ce qui prouve bien la continuité de $U_A(\cdot,\cdot)$ et donc l'existence d'un équilibre de Nash mixte. □

C.2 Preuve de la proposition 3.2

Démonstration. Nous reprenons la trame d'une démonstration dans [27] pour prouver une condition suffisante pour que l'équilibre de Nash soit unique. Prouver qu'un équilibre est unique dans notre cadre peut être fait en prouvant que pour toute partition d'équilibre $(s_m)_{m\in\{1,\dots,M\}}$:

$$\frac{\mathrm{d}s_m}{\mathrm{d}s_{m-1}} > 0 \text{ for } 2 \leq m \leq M-1. \tag{C.9}$$

En effet, soit $(s_m)_{m\in\{1,\dots,M\}}$ et $(s'_m)_{m\in\{1,\dots,M\}}$ deux partitions d'équilibre *distinctes*, c'est-à-dire $\exists\, m_0 \in \{1,\dots,M\}$ tel que $s_{m_0} \neq s'_{m_0}$. Sans perte de généralité, nous supposons de plus que $s_{m_0} < s'_{m_0}$. Alors, comme $(s_m)_{m\in\{1,\dots,M\}}$ et $(s'_m)_{m\in\{1,\dots,M\}}$ sont deux partitions d'équilibre, nous avons :

$$\frac{\mathrm{d}s_m}{\mathrm{d}s_{m-1}} \tag{C.10}$$

$$\frac{\mathrm{d}s'_m}{\mathrm{d}s'_{m-1}} \tag{C.11}$$

Ce qui implique

$$s_{m_0+i} < s'_{m_0+i}\ \forall\, i \in \{1,2,\dots,M-m_0+1\} \tag{C.12}$$

et en particulier $s_M < s'_M$, ce qui est une contradiction car par définition des partitions nous devons avoir $s_M = 1 = s'_M$.

Nous prouvons maintenant que (C.9) est vraie si $\frac{p_{\max}}{p_{\min}} \leq \sqrt{1+b}$.
Notons $M_m \triangleq \int_{s_{m-1}}^{s_m} p(s)\mathrm{d}s$ et

$$\begin{cases} v_m \triangleq \frac{\mathrm{d}s_m}{\mathrm{d}a_m} \\ w_m \triangleq \frac{\mathrm{d}a_m}{\mathrm{d}s_{m-1}} \end{cases} . \tag{C.13}$$

Considérons les caractérisations explicites des meilleures réponses des deux agents :

$$s_{m-1} = \frac{a_{m-1}+a_m}{2}\ , \tag{C.14}$$

et

$$a_m = \frac{\int_{s_{m-1}}^{s_m} sp(s)\mathrm{d}s}{(1+b)\int_{s_{m-1}}^{s_m} p(s)\mathrm{d}s}\ , \tag{C.15}$$

Nous pouvons alors prouver que ceci implique :

$$w_m + (v_{m-1})^{-1} = 2\ , \tag{C.16}$$

et

$$\begin{aligned}(1+b) =& \frac{\left[s_m p(s_m) v_m - s_{m-1}p(s_{m-1})(w_m)^{-1}\right] M_m}{M_m^2} \\ & \frac{-\left[p(s_m)v_m - p(s_{m-1})(w_m)^{-1}\right]\int_{s_{m-1}}^{s_m} sp(s)\mathrm{d}s}{M_m^2}\ ,\end{aligned} \tag{C.17}$$

Ce qui est équivalent à :

$$(1+b)M_m^2 = v_m \left[s_m M_m - \int_{s_{m-1}}^{s_m} sp(s)\mathrm{d}s \right] p(s_m) + w_m^{-1} \left[\int_{s_{m-1}}^{s_m} sp(s)\mathrm{d}s - s_{m-1}M_m \right] p(s_{m-1}) \ . \tag{C.18}$$

Observons maintenant que

$$v_1 = (1+b)\frac{M_1^2}{p(s_1)\left[s_1 M_1 - \int_0^{s_1} sp(s)\mathrm{d}s\right]} \tag{C.19}$$

$$\geq (1+b)\frac{p_{\min}^2 (s_1-0)^2}{p_{\max}^2 \frac{(s_1-0)^2}{2}} \tag{C.20}$$

$$= 2(1+b)\left(\frac{p_{\min}}{p_{\max}}\right)^2 \tag{C.21}$$

$$\geq 1 \tag{C.22}$$

La dernière inégalité est vraie parce que $\frac{p_{\max}}{p_{\min}} \leq \sqrt{2(1+b)}$.

De plus : $v_{m-1} \geq 1$ implique que $w_m \geq 1$ et

$$w_m \geq 1 \Rightarrow$$

$$v_m \geq \frac{(1+b)M_m^2 - \left[\int_{s_{m-1}}^{s_m} sp(s)\mathrm{d}s - s_{m-1}M_m\right] p(s_{m-1})}{\left[s_m M_m - \int_{s_{m-1}}^{s_m} sp(s)\mathrm{d}s\right] p(s_m)}. \tag{C.23}$$

Alors

$$\frac{(1+b)M_m^2 - \left[\int_{s_{m-1}}^{s_m} sp(s)\mathrm{d}s - s_{m-1}M_m\right] p(s_{m-1})}{\left[s_m M_m - \int_{s_{m-1}}^{s_m} sp(s)\mathrm{d}s\right] p(s_m)}$$

$$\geq \frac{(1+b)p_{\min}^2 (s_m - s_{m-1})^2 - p_{\max}^2 \frac{(s_m - s_{m-1})^2}{2}}{p_{\max}^2 \frac{(s_m - s_{m-1})^2}{2}} \tag{C.24}$$

$$= 2(1+b)\left(\frac{p_{\min}}{p_{\max}}\right)^2 - 1 \tag{C.25}$$

$$\geq 1 \tag{C.26}$$

si la condition suffisante $\frac{p_{\max}}{p_{\min}} \leq \sqrt{1+b}$ est vérifiée.

Pour résumé, nous avons

$$\forall m \geq 2, v_m \geq 1 \text{ et } w_m \geq 1 \ , \tag{C.27}$$

ce qui implique

$$\frac{\mathrm{d}s_m}{\mathrm{d}s_1} = \prod_{j=2}^{m} v_j w_j \geq 1 > 0. \tag{C.28}$$

et donc l'unicité de l'équilibre. □

C.3 Preuve de la proposition 3.3

Démonstration. Nous commençons par regarder la meilleure réponse du consommateur. Supposons fixées les actions de l'agrégateur $(a_m)_{m\in\{1,\dots,M\}}$, le consommateur doit résoudre :

$$\max_{\mathbf{s}=(s_m)_{m\in\{1,\dots,M\}}} -(s-a_m)^2\mathbb{1}_{s\in[s_{m-1},s_m]} \tag{C.29}$$

avec $s_0=0$ et $s_M=1$. Il est important de noter ici que maximiser l'équation ci-dessus implique que l'utilité bayésienne (définie en équation (3.9)) est maximisée.
Nous obtenons la meilleure réponse :

$$BR^C(\mathbf{a}) = (s_m)_{m\in\{0,\dots,M\}} \tag{C.30}$$

$$= \begin{cases} 0 \text{ for } m=0 \\ \frac{a_m+a_{m+1}}{2} \text{ for } 1\leq m\leq M-1 \\ 1 \text{ for } m=M \end{cases} \tag{C.31}$$

Dans ce cas, la condition du premier ordre s'écrit, pour tout $m\in\{1,\dots,M-1\}$:

$$\frac{\partial\sum_{j=1}^{M}\int_{s_{j-1}}^{s_j}(s-a_j)^2p(s)\mathrm{d}s}{\partial s_m}=0 \tag{C.32}$$

$$\Leftrightarrow \left[(s_m-a_m)^2-(s_m-a_{m+1})^2\right]p(s_m)=0 \tag{C.33}$$

$$\Leftrightarrow s_m=\frac{a_m+a_{m+1}}{2} \tag{C.34}$$

Alors la matrice Hessienne u_C est donnée par, $\forall\, 1\leq m_1,m_2\leq M-1$, $m_1\neq m_2$:

$$\begin{cases} \frac{\partial^2u_C}{\partial s_{m_1}^2}=2(a_{m_1+1}-a_{m_1})p(s_{m_1}) \\ \frac{\partial^2u_C}{\partial s_{m_1}\partial s_{m_2}}=0 \end{cases}, \tag{C.35}$$

qui est définie négative car $a_m<a_{m+1}$ pour tout m. Nous obtenons donc le même résultat que dans l'équation (3.37).

Deuxièmement, supposons fixée la partition de l'espace d'états $(s_m)_m$. L'agrégateur résout :

$$\max_{(a_m)_{m\in\{1,\dots,M\}}} -\sum_{m=1}^{M}\int_{s_{m-1}}^{s_m}\left[(s-a_m)^2+bc(a_m)\right]p(s)\mathrm{d}s \tag{C.36}$$

Ce qui est équivalent à :

$$\max_{a_m} -\int_{s_{m-1}}^{s_m}\left[(s-a_m)^2+bc(a_m)\right]p(s)\mathrm{d}s \tag{C.37}$$

pour tout $m\in\{1,\dots,M\}$.

En considérant chaque sous-problème d'optimisation de a_m, la condition du premier ordre est, pour tout m :

$$-\int_{s_{m-1}}^{s_m}[2(a-s)+bc'(a)]p(s)ds=0 \tag{C.38}$$

$$\Leftrightarrow 2a+bc'(a)=2\frac{\int_{s_{m-1}}^{s_m}sp(s)\mathrm{d}s}{\int_{s_{m-1}}^{s_m}p(s)\mathrm{d}s} \tag{C.39}$$

En utilisant la convexité et la croissance de c, cette équation fournit une unique solution plus petite que 1 en utilisant $2 + bc'(1) > 2 > 2s_m > 2\frac{\int_{s_{m-1}}^{s_m} sp(s)\mathrm{d}s}{\int_{s_{m-1}}^{s_m} p(s)\mathrm{d}s}$. Ce qui implique que :

$$BR^A(\mathbf{s}) = \left(\max(0, h^{-1}(2\frac{\int_{s_{m-1}}^{s_m} sp(s)\mathrm{d}s}{\int_{s_{m-1}}^{s_m} p(s)\mathrm{d}s})) \right)_{m\in\{1,\dots,M\}}. \tag{C.40}$$

□

C.4 Preuve de la proposition 3.4

Démonstration. Grâce à la proposition 3.3, nous avons :

$$a_m^{(n)} = \max(0, h^{-1}(2\frac{\int_{s_{m-1}^{(n-1)}}^{s_m^{(n-1)}} sp(s)\mathrm{d}s}{\int_{s_{m-1}^{(n-1)}}^{s_m^{(n-1)}} p(s)\mathrm{d}s})) \ . \tag{C.41}$$

Par le théorème des fonctions implicites, et en utilisant que :

- $\frac{\int_{s_{m-1}}^{s_m} sp(s)\mathrm{d}s}{\int_{s_{m-1}}^{s_m} p(s)\mathrm{d}s}$ est strictement croissant à la fois s_{m-1} et s_m ;
- h est strictement croissante pour une fonction convexe c et $b \geq 0$.

Nous obtenons que $a_m^{(n)}$ est croissant à la fois en $s_{m-1}^{(n-1)}$ et $s_m^{(n-1)}$ et que la croissance est stricte pour :

$$h^{-1}\left(2\frac{\int_{s_{m-1}^{(n-1)}}^{s_m^{(n-1)}} sp(s)\mathrm{d}s}{\int_{s_{m-1}^{(n-1)}}^{s_m^{(n-1)}} p(s)\mathrm{d}s}\right) > 0. \tag{C.42}$$

Alors, prenons à un point intérieur $\left(a_m^{(0)}\right)_{m\in\{1,\dots,M\}}$, c'est-à-dire $0 < a_1^{(0)} < a_2^{(0)} < \dots < a_M^{(0)} < 1$.
Pour tout $n \geq 1$, s'il existe $m_0^{(n)} \in \{1, \dots, M\}$ tel que $a_{m_0^{(n)}} = 0$, alors $a_{m_0^{(n)\prime}} = 0$ pour tout $m_0^{(n)\prime} < m_0^{(n)}$, comme $(a_m^{(n)})_m$ est croissant et supérieure à 0.

Alors les cellules qui dégénèrent sont celles les plus proches de zéro, comme la cellule numéro m dégénère si $a_{m-1} = a_m = a_{m+1}$. □

C.5 Preuve de la proposition 3.5

Démonstration. Observons que $c'(0) = 0$ implique $h^{-1}(0) = 0$ qui lui meme implique

$$\forall s' > 0,\ h^{-1}(2\frac{\int_0^{s'} sp(s)\mathrm{d}s}{\int_0^{s'} p(s)\mathrm{d}s}) > 0 \tag{C.43}$$

car $2\frac{\int_s^{s'} sp(s)\mathrm{d}s}{\int_0^{s'} p(s)\mathrm{d}s} > 0$ et h^{-1} est croissant. Ce qui implique le résultat. □

C.6 Preuve du théorème 3.7

Démonstration. Notons premièrement que, grâce à la proposition 3.5, la dynamique de meilleure réponse est bien définie à chaque itération.
Sachant que $\forall 1 \leq m \leq M$, a_m (respectivement $\forall 1 \leq m \leq M-1$, s_m) est croissant en s_{m-1} et s_m (respectivement a_m et a_{m+1}), le lemme 3.6 s'applique si la taille de la partition est constante au cours de la dynamique.

Donc, il est suffisant de montrer que la taille $M^{(n)}$ de la partition converge durant la dynamique de meilleure réponse pour obtenir le résultat. Comme $M^{(n)} \in \mathbb{N}$, cela impliquera effectivement que $M^{(n)}$ est constant à partir d'une certaine itération de la dynamique de meilleure réponse.

$(M^{(n)})_n$ est borné inférieurement par 1, car nous devons avoir au moins un intervalle grâce aux conditions $s_0 = 0$ et $s_M = 1$. De plus, c'est une suite non croissante. En effet, si la séquence d'actions dégénère, alors à l'itération suivante il est impossible de "créer" une nouvelle cellule, comme $s_m = \frac{a_m + a_{m+1}}{2}$. LA suite $(M^{(n)})_n$ est donc bien convergente. □

Annexe D

Démonstrations du chapitre 4

« If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is. »

– Remarque de John Von Neumann, en temps que conférencier, *First national meeting of the Association for Computing Machinery (1947).*

Dans cette annexe, nous présentons les démonstrations des théorèmes du Chapitre 4, ainsi qu'un calcul de la contrainte d'information très utile pour les simulations considérées.

D.1 Preuve du Théorème 4.7

Démonstration. **Paramètre** Choisissons un paramètre $\epsilon > 0$, et une distribution de probabilité $Q^\star_{s_0}$ telle que :

$$\max_{\substack{i\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} \left[\max_{k\in\mathcal{K}} H(A_{-i,k}|S_k(a_i), A_k) + \log_2 \chi_i - \kappa\, I_{Q^\star_{s_0}}(S_0; V_k) \right] < 0, \tag{D.1}$$

Fonction d'encodage f_0 Première étape : identification du déviateur.

L'encodeur procède au test statistique donné par l'équation (D.2) et définissons, pour une séquence d'actions donnée $a^n = (a_1^n, \dots, a_K^n) \in \mathcal{A}^n$, l'identité i du déviateur :

$$i \in \arg\min_{k\in\mathcal{K}} \sum_{a_{-k}\in\mathcal{A}_{-k}} \left| \frac{N(a_{-k}|a^n_{-k})}{n} - Q^\star_{-k}(a_{-k}) \right|. \tag{D.2}$$

Deuxième étape : comme la distribution de probabilité a_i^n est inconnue pour l'encodeur ainsi que le décodeur (il est donc impossible d'utiliser des arguments classiques de décodage typique), il y a besoin de l'encoder séparément. Pour cela, il est suffisant pour l'encodeur d'envoyer l'identité i du déviateur et la séquence de couleurs $c_i^n \in \Phi_i^n$ qui correspond à la séquence d'actions $a_i^n \in \mathcal{A}_i^n$ avec $c_i = \phi_i(a_i)$, et il utilisera un codage source-canal séparé comme expliqué ci-dessous.

Troisième étape : encodage des composantes $a_{-i}^n \in \mathcal{A}_{-i}^n$: comme dans [60], la séquence $a_{-i}^n \in \mathcal{A}_{-i}^n$ sera encodée en fonction de la transition de probabilité $\mathcal{T}$ et de la séquence $a_i^n \in \mathcal{A}_i^n$. Nous avons deux cas : soit le symbole $a_i \in \mathcal{A}_i^n$ a été utilisé avec une fréquence suffisamment grande, et dans ce cas les séquences de signaux $(s_k^n(a_i))_{k\in\mathcal{K}}$ sont suffisamment longues pour utiliser un schéma de codage source à la Tuncel [84] ; soit le symbole n'a pas été utilisé avec une fréquence suffisamment grande, et à nouveau les arguments classiques de typicité ne fonctionneront pas, comme pour l'identité du déviateur dans la deuxième étape.

L'encodeur sépare les séquences $s_k^n \in \mathcal{S}_k^n$ en sous-séquences $(s_k^{n_{a_i}})_{a_i\in\mathcal{A}_i}$ indexées par $a_i \in \mathcal{A}_i$ où $n_{a_i} = N(a_i|a_i^n)$ est le nombre d'occurrence de a_i dans a_i^n. La sous-séquence $s_k^{n_{a_i}} \in \mathcal{S}_k^{n_{a_i}}$ est de longueur $n_{a_i} \in \mathbb{N}$ et est tirée i.i.d. via la distribution de probabilité conjointe $Q_{-i}^\star \otimes \mathcal{T}_{a_i} \in \Delta(\mathcal{A}_{-i} \times \mathcal{S})$. L'encodeur évalue la partition $(\tilde{\mathcal{A}}_i, \tilde{\mathcal{A}}_i^c)$ des symboles $a_i \in \mathcal{A}_i$ définie comme suit :
Pour tout $\varepsilon > 0$, il existe $\bar{n}_1$ tel que la probabilité d'erreur de la partie conjointe de notre schéma de codage source-canal est bornée supérieurement par ε :

- $a_i \in \tilde{\mathcal{A}}_i$, si $N(a_i|a_i^n) = n_{a_i} \leq \bar{n}_1$ et alors la séquence $a_{-i}^{n_{a_i}} \in \mathcal{A}_{-i}^{n_{a_i}}$ ne peut être reconstruite par des arguments classiques de décodage typique, et sera donc encodée séparément, avec l'identité du déviateur i et la séquences de couleurs, comme nous l'expliquons ci-après.

- $a_i \in \tilde{\mathcal{A}}_i^c$ si $N(a_i|a_i^n) = n_{a_i} > \bar{n}_1$ et alors la séquence $a_{-i}^{n_{a_i}} \in \mathcal{A}_{-i}^{n_{a_i}}$ est encodée en utilisant un schéma de codage source-canal conjoint inspiré de [84].

Fragmentation Comme nous l'expliquions, notre schéma de codage est en deux parties. En effet, nous allons partitionner la séquence d'entrées du canal en deux sous-séquences $s_0^m = (s_0^{m_1}, s_0^{m_2})$ avec

$$m = m_1 + m_2, \tag{D.3}$$

m_1, la longueur de la séquence d'entrées du canal pour la partie conjointe de notre schéma, doit vérifier :

$$\max_{i\in\mathcal{K}} \; n \log \chi_i - m_1 \min_{k\in\mathcal{K}} I(S_0; V_k) \leq 0 \tag{D.4}$$

$$\Leftrightarrow \quad m_1 \geq \frac{n \log \chi_{i^*}}{I(S_0; V_{k^*})} \tag{D.5}$$

avec $i^* = \arg\max_{i\in\mathcal{K}} \log \chi_i$ et $k^* = \arg\max_{k\in\mathcal{K}} I(S_0; V_k)$.
m_2, la longueur de la séquence d'entrées du canal pour la partie séparée de notre schéma, doit vérifier :

$$\max_{\substack{i\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} \max_{k\in\mathcal{K}} \; nH(A_{-i,k}|S_k(a_i), A_k) - m_2 I(S_0; V_k) \leq 0 \tag{D.6}$$

$$\Leftrightarrow \quad m_2 \geq \frac{nH(A_{-i^{**},k^{**}}|S_{k^{**}}(a_{i^{**}}), A_{k^{**}})}{I(S_0; V_{k^{**}})} \tag{D.7}$$

avec i^{**} et k^{**} qui réalise le max dans :

$$\max_{\substack{i\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} \max_{k\in\mathcal{K}} nH(A_{-i,k}|S_k(a_i), A_k) - m_2 I(S_0; V_k). \tag{D.8}$$

Comme nous supposons que (D.1) est vérifiée, nous pouvons choisir m_1 et m_2 tels que (D.4) et (D.6) sont vraies.

Les choix de m_1 et m_2 sont cruciaux pour assurer un décodage unique comme nous le verrons plus loin.

Nous allons encoder l'identité i du déviateur, la séquence de couleurs correspondante c_i^n et la séquence $a_{-i}^{n_{a_i}}$ (pour $a_i \in \tilde{\mathcal{A}}_i$) dans $s_0^{m_1}$ en utilisant un codage source-canal séparé, et la séquence $a_{-i}^{n_{a_i}}$ (pour $a_i \in \tilde{\mathcal{A}}_i^c$) dans $s_0^{m_2}$ en utilisant un codage source-canal conjoint.

Encodage du déviateur i, des couleurs c_i^n et de $a_{-i}^{n_{a_i}}$ $(a_i \in \tilde{\mathcal{A}}_i)$
Premièrement, définissons

$$W_1 \triangleq |\mathcal{K}|.\chi_i^n.|\mathcal{A}_{-i}|^{\sum_{a_i \in \tilde{\mathcal{A}}_i} n_{a_i}}. \tag{D.9}$$

C'est le nombre de messages distincts dont nous avons besoin pour envoyer l'information aux décodeurs.

Construction du premier livre de codes : générer W_1 mots de codes canal $s_0^{m_1}(w_1)$, $w_1 \in \{1, 2, \ldots, W_1\}$ de longueur m_1 tels que chaque symbole est généré i.i.d. par rapport à la distribution de probabilité fixée $Q_{s_0}^\star$, et le révéler à l'encodeur et aux décodeurs.

L'encodeur, connaissant le message W_1, choisira $s_0^{m_1}(w_1)$, qui sera envoyé avec le mot de code du schéma de source-canal conjoint.

Encodage de $a_{-i}^{n_{a_i}}$ $(a_i \in \tilde{\mathcal{A}}_i^c)$ via un codage source-canal conjoint
Définissons

$$W_2 \triangleq 2^{n[H(A_{-i})+\epsilon]} \tag{D.10}$$

et générons deux livres de codes de longueurs distincte mais indexés de la même manière :

- Un livre des codes source, avec des mots de codes de longueur n générés i.i.d. via $Q_{a_{-i}}^\star$, et indexés par $a_{-i}^n(w_2)$ pour $1 \leq w_2 \leq W_2$.
- Un livre de codes canal, avec des mots de codes de longueur m_2 générés i.i.d. via $Q_{s_0}^\star$, et indexés par $s_0^{m_2}(w_2)$ pour $1 \leq w_2 \leq W_2$.

Les deux livres de codes sont révélés à l'encodeur et aux décodeurs.

Maintenant, connaissant a_{-i}^n, l'encodeur cherche le plus petit w_2 tel que $a_{-i}^n(w_2) = a_{-i}^n$, et choisit $s_0^{m_2}(w_2)$. Si un tel indice w_2 n'est pas trouvé, déclarer une erreur.
Finalement, l'encodeur concatène $s_0^{m_1}(w_1)$ and $s_0^{m_2}(w_2)$ et transmet

$$s_0^m \triangleq \big(s_0^{m_1}(w_1), s_0^{m_2}(w_2)\big). \tag{D.11}$$

Fonction de décodage g_k pour l'agent $k \in \mathcal{K}$: le décodeur, ayant observé s_k^n et v_k^m, va partitionner la séquence $v_k^m = (v_k^{m_1}, v_k^{m_2})$ (il est important de noter que cela est possible car dans ce schéma le décodeur doit connaitre la réalisation des livres de codes et donc connaitra la valeur des paramètres m_1 et m_2), et décodera en deux parties :

- Premièrement il doit décoder $w_1 \in \{1, 2, \dots, W_1\}$ pour pouvoir trouver l'indice i du déviateur (choisi par le test statistique), le couleurs $c_i^n \in \Phi_i^n$ et la séquence de symboles $a_{-i}^{n_{a_i}} \in \mathcal{A}_{-i}^{n_{a_i}}$ pour la transition de probabilité $\mathcal{T}_{a_i}$, contrôlé par le symbole $a_i \in \tilde{\mathcal{A}}_i$. Sachant $v_k^{m_1}$, le décodeur k essaye de trouver $\hat{w}_1$ tel que

$$(s_0^{m_1}(\hat{w}_1), v_k^{m_1}) \in \mathcal{T}_\epsilon^{m_1}(\mathcal{S}_0 \times \mathcal{V}_k) \tag{D.12}$$

  Si un ou plusieurs $\hat{w}_1$ existent, choisir le plus petit, sinon déclarer une erreur.
  Le décodage est possible si

$$\log_2 W_1 \leq m_1 \min_{k \in \mathcal{K}} I(S_0; V_k) - \epsilon \tag{D.13}$$

  ce qui est vrai par définition de W_1 (équation (D.9)) et grâce à l'inégalité sur m_1 (équation (D.4)).

- Après avoir trouvé i et c_i^n, et connaissant l'information adjacente $s_k^n \in \mathcal{S}_k$, le décodeur k décode une séquence unique $a_i^n \in \mathcal{A}_i^n$ pour la composante $i \in \mathcal{K}$.

Le décodeur $k \in \mathcal{K}$ connait maintenant avec probabilité proche de 1 la séquence d'actions complète $a_i^n \in \mathcal{A}_i^n$ et caractérise la partition $\tilde{\mathcal{A}}_i$ et ${\tilde{\mathcal{A}}_i}^c$ de l'ensemble des symboles $\mathcal{A}_i$.

- Pour la transition de probabilité $\mathcal{T}_{a_i}$, contrôlé par le symbole $a_i \in \tilde{\mathcal{A}}_i$, la séquence de symboles $a_{-i}^{n_{a_i}} \in \mathcal{A}_{-i}^{n_{a_i}}$ est décodé avec l'aide de l'indice $\hat{w}_1$ trouvé à la première étape.
- Pour la transition de probabilité $\mathcal{T}_{a_i}$, contrôlé par le symbole $a_i \in {\tilde{\mathcal{A}}_i}^c$, la séquence d'actions $a_{-i}^{n_{a_i}} \in \mathcal{A}_{-i}^{n_{a_i}}$ est décodée en utilisant un décodage typique. Le décodeur $k \in \mathcal{K}$ trouve un unique $\hat{w}_2$ qui vérifie simultanément

$$(a_{-i}^n(\hat{w}_2), (s_k^n(a_i), a_k^n)) \in \mathcal{T}_\epsilon^n(\mathcal{A}_{-i} \times (\mathcal{S}_k \times \mathcal{A}_k)) \tag{D.14}$$

  et

$$(s_0^{m_2}(\hat{w}_2), v_k^{m_2}) \in \mathcal{T}_\epsilon^{m_2}(\mathcal{S}_0 \times \mathcal{V}_k) \tag{D.15}$$

  Si $\hat{w}_2$ n'est pas unique ou n'existe pas, déclarer une erreur.

  Le décodage est possible si

$$\log_2 W_2 \leq \min_{\substack{k \in \mathcal{K}, \\ a_i \in \mathcal{A}_i}} nI(A_{-i}; S_k(a_i), A_k) + m_2 I(S_0; V_k) - \epsilon \tag{D.16}$$

  ce qui est vrai par définition de W_2 (équation (D.10)) et par l'inégalité sur m_2 (équation (D.6)).

Probabilité d'erreur • La première possibilité est que le test statistique renvoie un autre agent $j \neq i \in \mathcal{K}$. Cette probabilité d'erreur est bornée supérieurement (même preuve que [60]) et converge vers 0 avec n.

- Deuxième possibilité, le test statistique renvoie le bon déviateur $i \in \mathcal{K}$. A partir de maintenant, i est fixé. Dans ce cas, nous avons différentes erreurs, correspondant aux évènements suivants :
  - ➤ Il existe un mot de code dans le premier livre de codes canal qui n'est pas typique :

$$\mathcal{E}_1 = \{\exists\, w_1 \in \{1, 2, \dots, W_1\} \mid s_0^{m_1}(w_1) \notin \mathcal{T}_\epsilon^{m_1}(\mathcal{S}_0)\} \tag{D.17}$$

  - ➤ Il existe un mot de code dans le livre de codes source qui n'est pas typique (avec la source) :

$$\mathcal{E}_2 = \{\exists\, w_2 \in \{1, 2, \dots, W_2\} \mid a_{-i}^n(w_2) \notin \mathcal{T}_\epsilon^n(\mathcal{A}_{-i})\} \tag{D.18}$$

  - ➤ Il est important de savoir que, comme nous avons choisi de paramétrer les séquences source (dans le livre de codes), l'information adjacente dépend également de l'indice w_2. Donc, la troisième erreur est la suivante :
    Il existe un indice w_2 tel que le mot de code $a_{-i}^n(w_2)$ et l'information adjacente (pour l'agent k) $s_k^n(w_2)$ ne sont pas conjointement typiques :

$$\mathcal{E}_3(k) = \{\exists\, w_2 \in \{1, 2, \dots, W_2\} \mid (a_{-i}^n(w_2), (s_k^n(w_2), a_k^n)) \notin \mathcal{T}_\epsilon^n(\mathcal{A}_{-i} \times (\mathcal{S}_k \times \mathcal{A}_k))\} \tag{D.19}$$

  - ➤ Pour l'erreur précédente, la sortie du canal $v_k^{m_2}(w_2)$ dépend de l'entrée du canal $s_0^{m_2}(w_2)$, donc dépend de l'indice w_2 choisi. La quatrième probabilité d'erreur est :
    Il existe un indice w_2 tel que $s_0^{m_2}(w_2)$ et $v_k^{m_2}(w_2)$ ne sont pas conjointement typiques :

$$\mathcal{E}_4(k) = \{\exists\, w_2 \in \{1, 2, \dots, W_2\} \mid (s_0^{m_2}(w_2), v_k^{m_2}(w_2)) \notin \mathcal{T}_\epsilon^{m_2}(\mathcal{S}_0 \times \mathcal{V}_k)\} \tag{D.20}$$

  - ➤ Il existe un indice w_1 tel que $s_0^{m_1}(w_1)$ et $v_k^{m_1}(w_1)$ ne sont pas conjointement typiques :

$$\mathcal{E}_5(k) = \{\exists\, w_1 \in \{1, 2, \dots, W_1\} \mid (s_0^{m_1}(w_1), v_k^{m_1}(w_1)) \notin \mathcal{T}_\epsilon^{m_1}(\mathcal{S}_0 \times \mathcal{V}_k)\} \tag{D.21}$$

  - ➤ Il existe un indice $\hat{w}_2 \neq w_2$ tel que (D.14) et (D.15) sont vraies pour $s_k^n(w_2)$ et $v_k^{m_2}(w_2)$:

$$\begin{aligned}\mathcal{E}_6(k, i) = \{&\exists\, \hat{w}_2 \neq w_2 \text{ s.t. } (a_{-i}^n(\hat{w}_2), (s_k^n(w_2), a_k^n)) \in \mathcal{T}_\epsilon^n(\mathcal{A}_{-i} \times (\mathcal{S}_k \times \mathcal{A}_k)) \\ &\text{and } (s_0^{m_2}(\hat{w}_2), v_k^{m_2}(w_2)) \in \mathcal{T}_\epsilon^{m_2}(\mathcal{S}_0 \times \mathcal{V}_k)\}\end{aligned} \tag{D.22}$$

  - ➤ Il existe un indice $\hat{w}_1 \neq w_1$ tel que (D.12) est vrai pour $v_k^{m_1}(w_1)$:

$$\mathcal{E}_7(k) = \{\ \exists\, \hat{w}_1 \neq w_1 \text{ s.t. } (s_0^{m_1}(\hat{w}_1), v_k^{m_1}(w_1)) \in \mathcal{T}_\epsilon^{m_1}(\mathcal{S}_0 \times \mathcal{V}_k)\ \} \tag{D.23}$$

Par les propriétés des ensembles typiques, nous avons pour tout k :

$$\mathrm{P}(\mathcal{E}_1) \to 0 \text{ quand } m_1 \to \infty \tag{D.24}$$

$$\mathrm{P}(\mathcal{E}_2) \to 0 \text{ quand } n \to \infty \tag{D.25}$$

$$\mathrm{P}(\mathcal{E}_3(k)) \to 0 \text{ quand } n \to \infty \tag{D.26}$$

$$\mathrm{P}(\mathcal{E}_4(k)) \to 0 \text{ quand } m_2 \to \infty \tag{D.27}$$

$$\mathrm{P}(\mathcal{E}_5(k)) \to 0 \text{ quand } m_1 \to \infty \tag{D.28}$$

Nous bornons supérieurement ces probabilités par ϵ.

De plus, comme les deux conditions de typicité dans $\mathcal{E}_6$ sont indépendantes, la probabilité de cet évènement est égal au produit des probabilités correspondant aux deux conditions de typicité, donc, avec $\delta > 0$ suffisamment petit :

$$\mathrm{P}(\mathcal{E}_6(k,i)) = \mathrm{P}\Big(\exists \hat{w}_2 \neq w_2 \mid (a_{-i}^n(\hat{w}_2), (s_k^n(w_2), a_k^n)) \in \mathcal{T}_\epsilon^n(\mathcal{A}_{-i} \times (\mathcal{S}_k \times \mathcal{A}_k)) \text{ et } (s_0^{m_2}(\hat{w}_2), v_k^{m_2}(w_2)) \in \mathcal{T}_\epsilon^{m_2}(\mathcal{S}_0 \times \mathcal{V}_k)\Big) \tag{D.29}$$

$$\leq \sum_{\hat{w}_2 \neq w_2} \mathrm{P}\Big((a_{-i}^n(\hat{w}_2), (s_k^n(w_2), a_k^n)) \in \mathcal{T}_\epsilon^n(\mathcal{A}_{-i} \times (\mathcal{S}_k \times \mathcal{A}_k)) \text{ et } (s_0^{m_2}(\hat{w}_2), v_k^{m_2}(w_2)) \in \mathcal{T}_\epsilon^{m_2}(\mathcal{S}_0 \times \mathcal{V}_k)\Big) \tag{D.30}$$

$$= \sum_{\hat{w}_2 \neq w_2} \sum_{\substack{(a_{-i}^n, s_k^n, a_k^n) \in \mathcal{T}_\epsilon^n(\mathcal{A}_{-i} \times (\mathcal{S}_k \times \mathcal{A}_k)), \\ (s_0^{m_2}, v_k^{m_2}) \in \mathcal{T}_\epsilon^{m_2}(\mathcal{S}_0 \times \mathcal{V}_k)}} \mathrm{P}\Big((a_{-i}^n(\hat{w}_2), (s_k^n(w_2), a_k^n)) = (a_{-i}^n, s_k^n, a_k^n) \text{ et } (s_0^{m_2}(\hat{w}_2), v_k^{m_2}(w_2)) = (s_0^{m_2}, v_k^{m_2})\Big) \tag{D.31}$$

$$= \sum_{\hat{w}_2 \neq w_2} \sum_{\substack{(a_{-i}^n, s_k^n, a_k^n) \in \mathcal{T}_\epsilon^n(\mathcal{A}_{-i} \times (\mathcal{S}_k \times \mathcal{A}_k)), \\ (s_0^{m_2}, v_k^{m_2}) \in \mathcal{T}_\epsilon^{m_2}(\mathcal{S}_0 \times \mathcal{V}_k)}} \mathrm{P}\Big((a_{-i}^n(\hat{w}_2), (s_k^n(w_2), a_k^n)) = (a_{-i}^n, s_k^n, a_k^n)\Big) \times \mathrm{P}\Big((s_0^{m_2}(\hat{w}_2), v_k^{m_2}(w_2)) = (s_0^{m_2}, v_k^{m_2})\Big) \tag{D.32}$$

$$\leq \sum_{\hat{w}_2 \neq w_2} 2^{-n[I(A_{-i}; S_k(a_i), A_k) - \delta]} . 2^{-m_2[I(S_0; V_k) - \delta]} \tag{D.33}$$

$$\leq W_2 . 2^{-n[I(A_{-i}; S_k(a_i), A_k) - \delta]} . 2^{-m_2[I(S_0; V_k) - \delta]} \tag{D.34}$$

$$= 2^{\log_2 W_2 - nI(A_{-i}; S_k(a_i), A_k) + n\delta - m_2 I(S_0; V_k) + m_2 \delta} \tag{D.35}$$

$$= 2^{\log_2 W_2 - nI(A_{-i}; S_k(a_i), A_k) - m_2 I(S_0; V_k) + (n+m_2)\delta} \tag{D.36}$$

$$= 2^{nH(A_{-i,k} | S_k(a_i), A_k) - m_2 I(S_0; V_k) + (n+m_2)\delta + n\epsilon} \tag{D.37}$$

L'équation (D.29) vient de la définition de $\mathcal{E}_6$.
Les équations (D.30) et (D.31) viennent de l'inégalité de Boole.
L'équation (D.32) vient de l'indépendance des évènements.
Les équations (D.33) et (D.34) viennent des propriétés des ensembles typiques.

L'équation (D.35) vient de la contrainte d'information (D.16).
L'équation (D.37) vient du fait que $H(A_{-i,k}|S_k(a_i), A_k) = H(A_{-i}|S_k(a_i), A_k)$.

Pour la dernière probabilité d'erreur, nous avons :

$$\mathrm{P}(\mathcal{E}_7(k)) = \mathrm{P}\Big(\exists \hat{w}_1 \neq w_1 \mid (s_0^{m_1}(\hat{w}_1), v_k^{m_1}(w_1)) \in \mathcal{T}_\epsilon^{m_1}(\mathcal{S}_0 \times \mathcal{V}_k)\Big) \tag{D.38}$$

$$\leq \sum_{\hat{w}_1 \neq w_1} \mathrm{P}\Big((s_0^{m_1}(\hat{w}_1), v_k^{m_1}(w_1)) \in \mathcal{T}_\epsilon^{m_1}(\mathcal{S}_0 \times \mathcal{V}_k)\Big) \tag{D.39}$$

$$= \sum_{\hat{w}_1 \neq w_1} \sum_{(s_0^{m_1}, v_k^{m_1}) \in \mathcal{T}_\epsilon^{m_1}(\mathcal{S}_0 \times \mathcal{V}_k)} \mathrm{P}\Big((s_0^{m_1}(\hat{w}_1), v_k^{m_1}(w_1)) = (s_0^{m_1}, v_k^{m_1})\Big) \tag{D.40}$$

$$\leq \sum_{\hat{w}_1 \neq w_1} 2^{-m_1[I(S_0;V_k)-\delta]} \tag{D.41}$$

$$\leq W_1 . 2^{-m_1[I(S_0;V_k)-\delta]} \tag{D.42}$$

$$= 2^{\log_2 W_1 - m_1 I(S_0;V_k) + m_1\delta} \tag{D.43}$$

$$\leq 2^{-m_1 I(S_0;V_k) + n\log_2 \chi_i + \log_2 K + \bar{n}_1 |\tilde{\mathcal{A}}_i| \log_2 |\mathcal{A}_{-i}| + m_1\delta} \tag{D.44}$$

L'équation (D.38) vient de la définition de $\mathcal{E}_7$.
Les équations (D.39) et (D.40) viennent de l'inégalité de Boole.
L'équation (D.41) vient des propriétés des séquences typiques.

Ces inégalités sont vraies pour tout agent $k \in \mathcal{K}$ et tout déviateur $i \in \mathcal{K}$, nous pouvons donc borner la probabilité d'erreur totale comme suit :

$$\sum_{k\in\mathcal{K}} \max_{i\in\mathcal{K}} \max_{\mathcal{Q}_i \in \Delta(\mathcal{A}_i^\infty)} \mathrm{P}_{\mathcal{Q}_i}(a^n \neq g_k^{(m,n)}(s_k^n, v_k^m, a_k^n)) \tag{D.45}$$

$$\leq \sum_{k\in\mathcal{K}} \max_{\substack{i\in\mathcal{K}, \\ \mathcal{Q}_i \in \Delta(\mathcal{A}_i^\infty)}} \mathrm{P}(\mathcal{E}_1) + \mathrm{P}(\mathcal{E}_2) + \mathrm{P}(\mathcal{E}_3(k)) + \mathrm{P}(\mathcal{E}_4(k)) + \mathrm{P}(\mathcal{E}_5(k)) + \mathrm{P}(\mathcal{E}_6(k,i)) + \mathrm{P}(\mathcal{E}_7(k)) \tag{D.46}$$

$$\leq |\mathcal{K}| \Big[5\epsilon + 2^{\max_{i\in\mathcal{K}, a_i\in\mathcal{A}_i} \max_{k\in\mathcal{K}} \log_2 W_2 - nI(A_{-i};S_k(a_i),A_k) - m_2 I(S_0;V_k) + (n+m_2)\delta}$$

$$+ 2^{\max_{i\in\mathcal{K}, a_i\in\mathcal{A}_i} \max_{k\in\mathcal{K}} \log_2 W_1 - m_1 I(S_0;V_k) + m_1\delta} \Big] \tag{D.47}$$

$$\leq |\mathcal{K}| \Big[5\epsilon + 2^{(n+m_2)\delta} + 2^{m_1\delta} \Big] \tag{D.48}$$

$$\to_{\epsilon,\delta\to 0} 0 \tag{D.49}$$

L'équation (D.48) provient des contraintes d'information sur W_1 et W_2 (équations (D.13) et (D.16)). □

D.2 Preuve du Théorème 4.8

Nous commençons la encore par des cas simples, correspondant aux contraintes d'information (4.20), (4.21), (4.23), et enfin (4.26) :

1. (Sans information adjacente, sans robustesse)
   Nous supposons OPV : soit $\epsilon > 0$ et $\mathrm{P}_e^{(n,m)}(\lambda) \triangleq \sum_{k\in\mathcal{K}} \mathrm{P}(g_k^{(n,m)}(v_k^m, a_k^n) \neq a^n) < \epsilon$ pour une séquence de fonctions d'encodage $f_0^{(m,n)} : \mathcal{A}^n \mapsto \mathcal{S}_0^m$ et de fonctions de décodage $g_k^{(m,n)} : \mathcal{V}_k^m \times \mathcal{A}_k^n \mapsto \mathcal{A}^n$ avec un taux fixé à $\kappa = \frac{m}{n}$. Soit $\hat{A}_k^n \triangleq (\hat{A}_{1,k}^n, \hat{A}_{2,k}^n, \ldots, \hat{A}_{K,k}^n) \triangleq g_k^{(m,n)}(V_k^m)$.

   Pour tout $k \in \mathcal{K}$:

$$
\begin{aligned}
I(S_0; V_k) &\geq \frac{1}{m}\sum_{t=1}^{m} I(S_{0,t}; V_{k,t}) && \text{(D.50)}\\
&\geq \frac{1}{m} I(S_0^m; V_k^m) && \text{(D.51)}\\
&\geq \frac{1}{m} I(A^n; V_k^m) && \text{(D.52)}\\
&= \frac{1}{m}\big[H(A^n) - H(A^n|V_k^m)\big] && \text{(D.53)}\\
&\geq \frac{1}{m}\big[H(A^n) - H(A^n|\hat{A}_k^m)\big] && \text{(D.54)}\\
&= \frac{1}{\kappa} H(A) - \frac{1}{m} H(A^n|\hat{A}_k^n) && \text{(D.55)}\\
&\geq \frac{1}{\kappa} H(A) - \frac{1}{m}\big[1 + \mathrm{P}_e^{(n,m)}(\lambda) n \log|\mathcal{A}|\big] && \text{(D.56)}\\
&\geq \frac{1}{\kappa}\big[H(A_{-k}) - \frac{1}{n} - \mathrm{P}_e^{(n,m)}(\lambda)\log|\mathcal{A}|\big] && \text{(D.57)}\\
&\geq \frac{1}{\kappa}\big[H(A_{-k}) - \epsilon\big] && \text{(D.58)}
\end{aligned}
$$

   - L'équation (D.51) vient de l'hypothèse sans mémoire du canal de diffusion, et en utilisant le Lemme 7.9.2 de [26] nous avons $\sum_{t=1}^m I(S_{0,t}; V_{k,t}) \geq I(S_0^m; V_k^m)$.
   - L'équation (D.52) vient de l'inégalité de traitement des données A.12 avec la chaîne de Markov $A^n - S_0^m - V_k^m$.
   - L'équation (D.53) vient de la définition de l'information mutuelle A.7.
   - L'équation (D.54) vient de l'hypothèse $\hat{A}_k^n = g_k^{(m,n)}(V_k^m)$, qui implique que $A^n - V_k^n - \hat{A}^n$ est une chaine de Markov, et par l'inégalité de traitement des données que $H(A^n|V_k^m) \leq H(A^n|\hat{A}_k^m)$.
   - L'équation (D.55) est vraie parce que A^n est i.i.d.
   - L'équation (D.56) par l'inégalité de Fano [26].
   - L'équation (D.57) vient de $H(A_{-k}, A_k) \geq H(A_{-k})$.
   - L'équation (D.58) est vraie avec notre hypothèse OPV.

2. (Sans information adjacente, avec robustesse) :
   Nous supposons OVP : soit $\epsilon > 0$ et

$$
\mathrm{P}_e^{(n,m)}(\lambda) \triangleq \sum_{k\in\mathcal{K}} \max_{i\in\mathcal{K}} \max_{\mathcal{Q}_i \in \Delta(\mathcal{A}_i^\infty)} \mathrm{P}_{\mathcal{Q}_i}(g_k^{(n,m)}(v_k^m, a_k^n) \neq a^n) < \epsilon \qquad \text{(D.59)}
$$

pour une séquence de fonctions d'encodage $f_0^{(m,n)} : \mathcal{A}^n \mapsto \mathcal{S}_0^m$ et les fonctions de décodage $g_k^{(m,n)} : \mathcal{V}_k^m \times \mathcal{A}_k^n \mapsto \mathcal{A}^n$ avec un taux fixé à $\kappa = \frac{m}{n}$. Soit $\hat{A}_k^n \triangleq (\hat{A}_{1,k}^n, \hat{A}_{2,k}^n, \dots, \hat{A}_{K,k}^n) \triangleq g_k^{(m,n)}(V_k^m)$.

Pour un agent $k \in \mathcal{K}$ fixé, et pour un déviateur $i \in \mathcal{K}$ fixé, nous avons :

$$
\begin{aligned}
I(S_0; V_k) &\geq \frac{1}{m} \sum_{t=1}^{m} I(S_{0,t}; V_{k,t}) && \text{(D.60)}\\
&\geq \frac{1}{m} I(S_0^m; V_k^m) && \text{(D.61)}\\
&\geq \frac{1}{m} I(A^n; V_k^m) && \text{(D.62)}\\
&= \frac{1}{m} \big[H(A^n) - H(A^n | V_k^m) \big] && \text{(D.63)}\\
&= \frac{1}{m} \big[H(A^n) - H(A^n | \hat{A}_k^m) \big] && \text{(D.64)}\\
&\geq \frac{1}{\kappa} H(A) - \frac{1}{m} H(A^n | \hat{A}_k^n) && \text{(D.65)}\\
&= \frac{1}{\kappa} \big[H(A_{-i}) + H(A_i | A_{-i}) \big] - \frac{1}{m} H(A^n | \hat{A}_k^n) && \text{(D.66)}\\
&\geq \frac{1}{\kappa} \big[H(A_{-i}) + H(A_i | A_{-i}) \big] - \frac{1}{m} \big[1 + \mathrm{P}_e^{(n,m)}(\lambda) n \log |\mathcal{A}| \big] && \text{(D.67)}\\
&\geq \frac{1}{\kappa} \big[H(A_{-i,k}) + H(A_i | A_{-i}) - \frac{1}{n} - \mathrm{P}_e^{(n,m)}(\lambda) \log |\mathcal{A}| \big] && \text{(D.68)}\\
&\geq \frac{1}{\kappa} \big[H(A_{-i,k}) - \epsilon + H(A_i | A_{-i}) \big] && \text{(D.69)}\\
&= \frac{1}{\kappa} \big[H(A_{-i,k}) - \epsilon + H(A_i) \big] && \text{(D.70)}
\end{aligned}
$$

- L'équation (D.61) vient de l'hypothèse sans mémoire sur le canal de diffusion, et en utilisant le Lemme 7.9.2 of [26] nous avons $\sum_{t=1}^{m} I(S_{0,t}; V_{k,t}) \geq I(S_0^m; V_k^m)$.
- L'équation (D.62) vient de l'inégalité de traitement des données A.12 et de la chaîne de Markov $A^n - S_0^m - V_k^m$.
- L'équation (D.63) vient de la définition de l'information mutuelle A.7.
- L'équation (D.64) vient de l'hypothèse $\hat{A}_k^n = g_k^{(m,n)}(V_k^m)$, qui implique $A^n - V_k^n - \hat{A}^n$ est une chaîne de Markov, et par l'inégalité de traitement des données : $H(A^n | V_k^m) \leq H(A^n | \hat{A}_k^m)$.
- L'équation (D.65) est vraie parce que A^n est i.i.d.
- L'équation (D.66) est la règle de chaîne pour l'entropie A.6.
- L'équation (D.67) par l'inégalité de Fano [26].
- L'équation (D.69) est vraie grâce à l'hypothèse OVP.
- L'équation (D.70) est vraie si les agents jouent des stratégies mixtes, c'est-à-dire que $(A_i)_{i \in \mathcal{K}}$ sont indépendants ;
- L'équation (D.66) est vraie comme la distribution de probabilité sur $\mathcal{A}_i$ est inconnue, nous devons coder au taux $\log_2 |\mathcal{A}_i|$.

Ceci est vrai pour tout k et i, donc nous pouvons prendre le maximum sur $k, i \in \mathcal{K}$. De plus, comme la distribution de probabilité sur $\mathcal{A}_i$ est inconnue, nous devons prendre le maximum sur toutes les distributions de probabilité $Q_i \in \mathcal{A}_i$:

$$\max_{\substack{i\in\mathcal{K},\\ Q_i\in\Delta(\mathcal{A}_i)}} \max_{k\in\mathcal{K}} \left[H(A_{-i,k}) + H_{Q_i}(A_i) - \kappa I(S_0;V_k)\right] \leq \epsilon \tag{D.71}$$

$$\Leftrightarrow \max_{i\in\mathcal{K}} \max_{k\in\mathcal{K}} \left[H(A_{-i,k}) + \max_{Q_i\in\Delta(\mathcal{A}_i)} H_{Q_i}(A_i) - \kappa I(S_0;V_k)\right] \leq \epsilon \tag{D.72}$$

$$\Leftrightarrow \max_{i\in\mathcal{K}} \max_{k\in\mathcal{K}} \left[H(A_{-i,k}) + \log_2|\mathcal{A}_i| - \kappa I(S_0;V_k)\right] \leq \epsilon \tag{D.73}$$

L'équation (D.73) vient du fait que : $\max_{Q_i\in\Delta(\mathcal{A}_i)} H_{Q_i}(A_i) = \log_2|\mathcal{A}_i|$. QED

3. (Avec information adjacente, sans robustesse)
C'est la preuve réciproque de [84].

Preuve de Tuncel. Soit $\epsilon > 0$ et $\mathrm{P}_e^{(n,m)}(\lambda) \triangleq \sum_{k\in\mathcal{K}} \mathrm{P}(g_k^{(n,m)}(s_k^n, v_k^m, a_k^n) \neq a^n) < \epsilon$ pour une séquence de fonctions d'encodage $f_0^{(m,n)} : \mathcal{A}^n \mapsto \mathcal{S}_0^m$ et de fonctions de décodage $g_k^{(m,n)} : \mathcal{V}_k^m \times \mathcal{S}_k^n \times \mathcal{A}_k^n \mapsto \mathcal{A}^n$ avec un taux fixé à $\kappa = \frac{m}{n}$. Soit $\hat{A}_k^n \triangleq (\hat{A}_{1,k}^n, \hat{A}_{2,k}^n, \ldots, \hat{A}_{K,k}^n) \triangleq g_k^{(m,n)}(V_k^m, S_k^n)$.

$$\begin{aligned}
I(S_0;V_k) &\geq \frac{1}{m}\sum_{t=1}^{m} I(S_{0,t};V_{k,t}) && \text{(D.74)}\\
&\geq \frac{1}{m} I(S_0^m;V_k^m) && \text{(D.75)}\\
&\geq \frac{1}{m} I(A^n;V_k^m) && \text{(D.76)}\\
&\geq \frac{1}{m} I(A^n;V_k^m|S_k^n) && \text{(D.77)}\\
&= \frac{1}{m}\left[H(A^n|S_k^n) - H(A^n|S_k^n,V_k^m)\right] && \text{(D.78)}\\
&\geq \frac{1}{m}\left[H(A^n|S_k^n) - H(A^n|\hat{A}_k^n)\right] && \text{(D.79)}\\
&\geq \frac{1}{\kappa} H(A|S_k) - \frac{1}{m} H(A^n|\hat{A}_k^n) && \text{(D.80)}\\
&\geq \frac{1}{\kappa} H(A|S_k) - \frac{1}{m}\left[1 + \mathrm{P}_e^{(n,m)}(\lambda) n \log|\mathcal{A}|\right] && \text{(D.81)}\\
&= \frac{1}{\kappa}\left[H(A|S_k) - \frac{1}{n} - \mathrm{P}_e^{(n,m)}(\lambda)\log|\mathcal{A}|\right] && \text{(D.82)}\\
&\geq \frac{1}{\kappa}\left[H(A|S_k) - \epsilon\right] && \text{(D.83)}
\end{aligned}$$

- L'équation (D.75) vient de l'hypothèse sans mémoire sur le canal de diffusion, et en utilisant le Lemme 7.9.2 of [26] nous avons $\sum_{t=1}^{m} I(S_{0,t};V_{k,t}) \geq I(S_0^m;V_k^m)$.
- L'équation (D.76) vient de l'inégalité de traitement des données A.12 et de la chaîne de Markov $A^n - S_0^m - V_k^m$.

- L'équation (D.77) vient de la chaîne de Markov $S_k^n - A^n - V_k^m$.
- L'équation (D.78) vient de la définition de l'information mutuelle conditionnelle A.10.
- L'équation (D.79) vient de l'hypothèse $\hat{A}_k^n = g_k^{(m,n)}(V_k^m, S_k^n)$, qui implique que $A^n - V_k^n, S_k^n - \hat{A}^n$ est une chaîne de Markov, et par l'inégalité de traitement des données que $H(A^n|V_k^m) \leq H(A^n|\hat{A}_k^m)$.
- L'équation (D.81) par l'inégalité de Fano [26].
- L'équation (D.83) est vrai par l'hypothèse OVP.

□

4. (Avec information adjacente, avec robustesse) : preuve du théorème 4.8.
Soit $\epsilon > 0$ et

$$\mathrm{P}_e^{(n,m)}(\lambda) \triangleq \sum_{k\in\mathcal{K}} \max_{i\in\mathcal{K}} \max_{\mathcal{Q}_i\in\Delta(\mathcal{A}_i^\infty)} \mathrm{P}_{\mathcal{Q}_i}(g_k^{(n,m)}(s_k^n, v_k^m, a_k^n) \neq a^n) < \epsilon \tag{D.84}$$

pour une séquence de fonctions d'encodage $f_0^{(m,n)} : \mathcal{A}^n \mapsto \mathcal{S}_0^m$ et de fonctions de décodage $g_k^{(m,n)} : \mathcal{V}_k^m \times \mathcal{S}_k^n \times \mathcal{A}_k^n \mapsto \mathcal{A}^n$ avec un taux fixé à $\kappa = \frac{m}{n}$. Soit $\hat{A}_k^n \triangleq (\hat{A}_{1,k}^n, \hat{A}_{2,k}^n, \ldots, \hat{A}_{K,k}^n) \triangleq g_k^{(m,n)}(V_k^m, S_k^n)$.

Dans notre cadre, (D.75),(D.76) et (D.77) sont toujours vraies. Nous pouvons commencer après l'équation (D.77) :

$$\begin{aligned}
(\text{D.77}) \quad \frac{1}{m} I(A^n, V_k^m|S_k^n) &= \frac{1}{m}\Big[H(A^n|S_k^n) + H(A^n|V_k^m, S_k^n)\Big] && \text{(D.85)}\\
&= \frac{1}{m}\Big[H(A_i^n|S_k^n) + H(A_{-i}^n|A_i^n, S_k^n) - H(A^n|V_k^m, S_k^n)\Big] && \text{(D.86)}\\
&= \frac{1}{\kappa}\Big[H(A_i|S_k) + H(A_{-i}|A_i, S_k)\Big] - \frac{1}{m} H(A^n|V_k^n, S_k^n) && \text{(D.87)}\\
&\geq \frac{1}{\kappa}\Big[H(A_i|S_k) + H(A_{-i}|A_i, S_k)\Big] - \frac{1}{m} H(A^n|\hat{A}_k^n) && \text{(D.88)}\\
&\geq \frac{1}{\kappa}\Big[H(A_i|S_k) + H(A_{-i}|A_i, S_k)\Big] - \frac{1}{m}\big[1 + \mathrm{P}_k^{(m,n)} n \log|\mathcal{A}|\big] && \text{(D.89)}\\
&\geq \frac{1}{\kappa}\Big[H(A_i|S_k) + H(A_{-i,k}|A_i, S_k)\Big] - \epsilon && \text{(D.90)}\\
&= \frac{1}{\kappa}\Big[H(A_i|S_k) + H(A_{-i,k}|S_k)\Big] - \epsilon && \text{(D.91)}\\
&\geq \frac{1}{\kappa}\Big[H(A_i|S_k) + H(A_{-i,k}|S_k, A_k)\Big] - \epsilon && \text{(D.92)}
\end{aligned}$$

- L'équation (D.86) vient de la règle de chaîne pour l'entropie A.6.
- L'équation (D.88) vient de l'hypothèse $\hat{A}_k^n = g_k^{(m,n)}(V_k^m, S_k^n)$, qui implique que $A^n - V_k^n, S_k^n - \hat{A}^n$ est une chaîne de Markov, et par l'inégalité de traitement des données que $H(A^n|V_k^m, S_k^n) \leq H(A^n|\hat{A}_k^m)$.

- L'équation (D.89) par l'inégalité de Fano [26].
- L'équation (D.90) vient de notre hypothèse OVP, et du fait que $H(A_{-i}|A_i, S_k) \geq H(A_{-i,k}|A_i, S_k)$.
- L'équation (D.91) vient du fait que S_k est une fonction de A_i.
- L'équation (D.92) vient du fait que conditionner réduit l'entropie A.9.

Ceci est vrai pour tout k et i, donc nous pouvons prendre la maximum sur $k, i \in \mathcal{K}$. De plus, comme la distribution de probabilité sur $\mathcal{A}_i$ est inconnue, nous devons considérer la pire action en terme d'efficacité de codage $a_i \in \mathcal{A}_i$ du déviateur i :

$$\max_{\substack{i\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} \max_{k\in\mathcal{K}} \left[H(A_{-i,k}|A_k, S_k(a_i)) + H(A_i|S_k(a_i)) - \kappa I(S_0; V_k)\right] \leq \epsilon \tag{D.93}$$

$$\Leftrightarrow \max_{\substack{i\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} \max_{k\in\mathcal{K}} \left[H(A_{-i,k}|A_k, S_k(a_i)) + H(C_i|S_k(a_i)) - \kappa I(S_0; V_k)\right] \leq \epsilon \tag{D.94}$$

L'équation (D.94) est vraie car :
Soit C_i la variable aléatoire des couleurs correspondant à la séquence d'action de l'agent i :

$$H(A_i|S_k(a_i)) = H(A_i|S_k(a_i), C_i) + H(C_i|S_k(a_i)) - H(C_i|A_i, S_k(a_i)) \tag{D.95}$$

$$= H(C_i|S_k(a_i)) \tag{D.96}$$

parce que :

- Le coloriage est une fonction des actions, donc $H(C_i|A_i) = H(C_i|A_i, S_k(a_i)) = 0$;
- Le coloriage et les signaux permettent de retrouver les actions : $H(A_i|S_k(a_i), C_i) = 0$.

De plus, pour un décodeur k, nous considérons la pire déviation de l'agent i en terme d'efficacité de codage, ce qui se traduit par une incertitude maximale sur la variable aléatoire C_i des couleurs lorsque la réalisation de la variable des signaux S_k est observée. L'entropie conditionnelle vaut donc $H(C_i|S_k(a_i)) = \log_2 \chi_i$. Dès lors :

$$\max_{\substack{i\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} \max_{k\in\mathcal{K}} \left[H(A_{-i,k}|A_k, S_k(a_i)) + H(C_i|S_k(a_i)) - \kappa I(S_0; V_k)\right] \leq \epsilon \tag{D.97}$$

$$\Leftrightarrow \max_{\substack{i\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} \max_{k\in\mathcal{K}} \left[H(A_{-i,k}|A_k, S_k(a_i)) + \log_2 \chi_i - \kappa I(S_0; V_k)\right] \leq \epsilon. \qquad \square \tag{D.98}$$

D.3 Écriture de la contrainte d'information pour l'application considérée

Pour le problème exposé en Section 4.4, nous pouvons réécrire la contrainte d'information comme suit.
Soit $Q^\star_{S_0} \in \Delta(\mathcal{S}_0)$. Nous avons :

$$I_{Q^\star_{S_0}}(S_0; V_k) = \sum_{s_0\in\mathcal{S}_0} Q^\star(s_0) \log_2 \frac{1}{Q^\star(s_0)} - \sum_{\substack{s_0\in\mathcal{S}_0,\\ v_k\in\mathcal{V}_k}} Q^\star(s_0)\mathcal{T}_C(v_k|s_0) \log_2 \frac{\sum_{\tilde{s}_0\in\mathcal{S}_0} Q^\star(\tilde{s}_0)\mathcal{T}_C(v_k|\tilde{s}_0)}{Q^\star(s_0)\mathcal{T}_C(v_k|s_0)} \tag{D.99}$$

Nous avons alors :

$$\max_{\substack{i\in\mathcal{K},\\ a_i\in\mathcal{A}_i}} \left[\max_{k\in\mathcal{K}} H(A_{-i,k}|S_k(a_i),A_k) + \log_2\chi_i - I_{Q^\star_{S_0}}(S_0;V_k)\right] < 0$$

$$\overset{(a)}{\Leftrightarrow} \quad \max_{k\in\mathcal{K}} H(A_{-k}|S_k(a_k),A_k) + \log_2\chi_k - I_{Q^\star_{S_0}}(S_0;V_k) < 0$$

$$\overset{(b)}{\Leftrightarrow} \quad \max\begin{bmatrix} H(a_2|s_1) + \log_2\chi_1 - I_{Q^\star_{S_0}}(S_0;V_1), \\ H(a_1|s_2) + \log_2\chi_2 - I_{Q^\star_{S_0}}(S_0;V_2) \end{bmatrix} < 0$$

$$\overset{(c)}{\Leftrightarrow} \quad \max\begin{bmatrix} \sum_{\substack{a_2\in\mathcal{A}_2,\\ s_1\in\mathcal{S}_1}} Q^\star(a_2)\urcorner(s_1|a_2)\log_2\frac{\sum_{\tilde{a}_2\in\mathcal{A}_2}Q^\star(\tilde{a}_2)\urcorner(s_1|\tilde{a}_2)}{Q^\star(a_2)\urcorner(s_1|a_2)} + \log_2\chi_1 - I_{Q^\star_{S_0}}(S_0;V_1), \\ \sum_{\substack{a_1\in\mathcal{A}_1,\\ s_2\in\mathcal{S}_2}} Q^\star(a_1)\urcorner(s_2|a_1)\log_2\frac{\sum_{\tilde{a}_1\in\mathcal{A}_1}Q^\star(\tilde{a}_1)\urcorner(s_2|\tilde{a}_1)}{Q^\star(a_1)\urcorner(s_2|a_1)} + \log_2\chi_2 - I_{Q^\star_{S_0}}(S_0;V_2) \end{bmatrix} < 0$$

$$\overset{(d)}{\Leftrightarrow} \quad \max\begin{bmatrix} & Q^\star(a_2)(1-\frac{\delta}{2})\cdot\log_2\left(\frac{Q^\star(a_2)(1-\frac{\delta}{2})+Q^\star(a_2')\frac{\delta}{2}}{Q^\star(a_2)(1-\frac{\delta}{2})}\right) \\ + & Q^\star(a_2')\frac{\delta}{2}\cdot\log_2\left(\frac{Q^\star(a_2)(1-\frac{\delta}{2})+Q^\star(a_2')\frac{\delta}{2}}{Q^\star(a_2')\frac{\delta}{2}}\right) \\ + & Q^\star(a_2)\frac{\delta}{2}\cdot\log_2\left(\frac{Q^\star(a_2)\frac{\delta}{2}+Q^\star(a_2')(1-\frac{\delta}{2})}{Q^\star(a_2)\frac{\delta}{2}}\right) \\ + & Q^\star(a_2')(1-\frac{\delta}{2})\cdot\log_2\left(\frac{Q^\star(a_2)\frac{\delta}{2}+Q^\star(a_2')(1-\frac{\delta}{2})}{Q^\star(a_2')(1-\frac{\delta}{2})}\right) - I_{Q^\star_{S_0}}(S_0;V_1), \\ & Q^\star(a_1)(1-\frac{\delta}{2})\cdot\log_2\left(\frac{Q^\star(a_1)(1-\frac{\delta}{2})+Q^\star(a_1')\frac{\delta}{2}}{Q^\star(a_1)(1-\frac{\delta}{2})}\right) \\ + & Q^\star(a_1')\frac{\delta}{2}\cdot\log_2\left(\frac{Q^\star(a_1)(1-\frac{\delta}{2})+Q^\star(a_1')\frac{\delta}{2}}{Q^\star(a_1')\frac{\delta}{2}}\right) \\ + & Q^\star(a_1)\frac{\delta}{2}\cdot\log_2\left(\frac{Q^\star(a_1)\frac{\delta}{2}+Q^\star(a_1')(1-\frac{\delta}{2})}{Q^\star(a_1)\frac{\delta}{2}}\right) \\ + & Q^\star(a_1')(1-\frac{\delta}{2})\cdot\log_2\left(\frac{Q^\star(a_1)\frac{\delta}{2}+Q^\star(a_1')(1-\frac{\delta}{2})}{Q^\star(a_1')(1-\frac{\delta}{2})}\right) - I_{Q^\star_{S_0}}(S_0;V_1) \end{bmatrix} < -1$$

où

- (a) vient du fait que si $i \neq k$, alors $a_{-i,k} = \emptyset$ donc nous avons seulement à considérer le cas où $i = k$;
- (b) vient du fait que a_2 et s_1 sont indépendants de a_1 et alors l'entropie $H(a_2|s_1,a_1)$ se réduit à $H(a_2|s_1)$. En utilisant le même argument, $H(a_1|s_2,a_2)$ se réduit à $H(a_1|s_2)$;
- (c) vient de la définition de l'entropie conditionnelle ;
- (d) vient du fait que les nombres chromatiques des graphes $\mathcal{G}_1$, $\mathcal{G}_2$ des deux agents sont égaux à $\chi_1 = \chi_2 = 2$ dès lors que $\delta > 0$.

Annexe E

Publications

E.1 Revues internationales

E.1.1 IEEE-TIT 2014

Coordination in distributed networks via coded actions with application to power control

Benjamin Larrousse, Samson Lasaulce, and Matthieu Bloch
{larrousse, lasaulce}@lss.supelec.fr, matthieu.bloch@ece.gatech.edu

Abstract

This paper investigates the problem of coordinating several agents through their actions. Although the methodology applies to general scenarios, the present work focuses on a situation with an asymmetric observation structure that only involves two agents. More precisely, one of the agents knows the past, present, and future realizations of a state that affects the common payoff function of the agents; in contrast, the second agent is assumed either to know the past realizations of the state or to have no knowledge of it. In both cases, the second agent has access to some observations of the first agent's actions, which enables the two agents to coordinate. These scenarios are applied to the problem of distributed power control; the key idea is that a transmitter may embed information about the channel state into its transmit power levels so that an observation of these levels, e.g., the signal-to-interference plus noise ratio, allows the other transmitter to coordinate its power levels. The main contributions of this paper are twofold. First, we provide a characterization of the set of feasible average payoffs when the agents repeatedly take long sequences of actions and the realizations of the state are i.i.d.. Second, we exploit these results in the context of distributed power control and introduce the concept of coded power control. We carry out an extensive numerical analysis of the benefits of coded power control over alternative power control policies, and highlight a simple yet non-trivial example of a power-control code.

Index Terms

channels with state; coding theorems; coordination; distributed power control; distributed resource allocation; game theory; information constraints; optimization.

I. Introduction

The main technical problem studied in this paper is the following. Given an integer $T \geq 1$, three discrete alphabets $\mathcal{X}_0$, $\mathcal{X}_1$, $\mathcal{X}_2$, and a *payoff* function $w : \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \to \mathbb{R}$, one wants to maximize the average payoff

$$W_T(x_0^T, x_1^T, x_2^T) \triangleq \frac{1}{T} \sum_{t=1}^{T} w(x_{0,t}, x_{1,t}, x_{2,t}) \tag{1}$$

with respect to the sequences $x_1^T \triangleq (x_{1,1}, \cdots, x_{1,T}) \in \mathcal{X}_1^T$ and $x_2^T \triangleq (x_{2,1}, \cdots, x_{2,T}) \in \mathcal{X}_2^T$ given the knowledge of $x_0^T \triangleq (x_{0,1}, \cdots, x_{0,T}) \in \mathcal{X}_0^T$. Without further restrictions and with instantaneous knowledge of $x_{0,t}$, solving this optimization problem merely consists in finding one of the optimal pairs of variables $(x_{1,t}^\star, x_{2,t}^\star)$ for every t. The corresponding maximum value[1] of W_T is then

$$W_T^\star = \frac{1}{T} \sum_{t=1}^{T} \max_{x_1, x_2} w(x_{0,t}, x_1, x_2). \tag{2}$$

In this paper, we introduce the additional restriction that the variable x_2 cannot be controlled or optimized directly. As formally described in Section II, the variable x_2 is required to result from imperfect observations of x_0 through x_1, which induces an *information constraint* in the aforementioned optimization problem. One of the main results of this paper in Section III is to precisely characterize this constraint for large T when x_0^T consists of independent identically distributed (i.i.d.) realizations of a given random variable X_0.

The problem at hand is a special case of a *distributed optimization problem*, in which K agents[2] connected via a given observation structure have the common objective of maximizing the average payoff W_T for large T. In this general framework, the variable x_k with $k \in \{1, \cdots, K\}$ is called the *action* of Agent k and represents the only variable under its control. The variable x_0 is outside of the agents' control and typically represents the realization of a random state. The observation structure defines how the agents interact through observations of the random state and of each other's actions. The average payoff then measures the degree of *coordination* between the agents, under the observation constraints of the actions imposed by the

This paper was presented in part at the 2013 IEEE International Symposium on Information Theory [1], the 2013 International Workshop on Stochastic Methods in Game Theory, and the 12th IEEE International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks in 2014 [2]. The work of B. Larrousse and S. Lasaulce was supported in part by the French agency ANR through the project LIMICOS - ANR-12-BS03-0005. The work of M. Bloch was supported in part by NSF under award CCF-1320304.

[1] In this paper, this scenario will be referred to as the costless communication case. In Section V, the corresponding power control scenario will be called costless communication power control (CCPC).

[2] In other disciplines such as computer science, control, or economics, agents are sometimes called nodes, controllers, or decision-makers.

observation structure. As a concrete example, we apply this framework to power control in Section V, in which x_0 represents the global channel state information (CSI) and x_k the power level of Transmitter k.

A central question in this general framework is to characterize the possible values of the average payoff W_T when the agents interact many times, i.e., when T is large. Since answering this question in its full generality still appears out of reach, the present paper tackles a special case with $K = 2$ agents. Specifically, we assume that Agent 1 has perfect knowledge of the past, current, and future realizations of the state sequence x_0^T, while Agent 2 obtains imperfect observations of Agent 1's actions and possesses either strictly causal or no knowledge of the realizations of the state. In spite of these restricting assumptions, it is possible to extract valuable concepts and insights of practical interest from the present work, which can be applied to the general case of $K \geq 2$ agents and arbitrary observation structures.

A. Related work

In most of the literature on coordination among agents, which includes classical team decision problems [3], the typical assumption is that agents have access to *dedicated channels* to coordinate their actions. Specifically, these dedicated channels allow the agents to signal or communicate with each other without affecting the objective or payoff function. The works on coordination with dedicated channels that relates most closely to the present work are [4], [5]. Therein, the authors introduce the notions of *empirical coordination* and *strong coordination* to measure the ability of agents to coordinate their actions in a network with noiseless dedicated channels. Empirical distribution measures an average coordination behavior over time and requires the joint empirical distribution of the actions to approach a target distribution asymptotically in variational distance; this notion relates to earlier work on the communication of probability distribution [6], and is analyzed using tools from rate-distortion theory. Strong coordination is more stringent and requires that the distribution of *sequences* of actions be asymptotically indistinguishable from sequences of actions drawn according to a target distribution, again in terms of variational distance; this notion relates to the concept of channel resolvability [7]. In both cases, the goal is to establish the *coordination capacity*, which relates the achievable joint distributions of actions to the fixed rate constraints on the noiseless dedicated channels. The results of [4], [5] have been extended to a variety of networks with dedicated channels [8], [9], [10], [11], and optimal codes have been designed for specific settings [12], [13].

In contrast, much less is known about the problem of coordination via the actions of agents in the absence of dedicated channels, which is the main topic of the present work. The most closely related work is [14], in which the authors analyze the problem of characterizing the set of possible average payoffs in the presence of two agents, assuming that each agent can perfectly monitor the actions of the other agent; the authors establish the set of *implementable distributions*, which are the achievable empirical joint distributions of the actions under the assumed observation structure. In particular, this set is characterized by an information constraint that captures the observation structure between the agents. While [14] largely relies on combinatorial arguments, [15] provides a more traditional information-theoretic approach of coordination via actions under the name of *implicit communication*.

To the best of our knowledge, the present work is the first to apply the framework of coordination through actions to the problem of distributed resource allocation in wireless networks. More specifically, the application developed in this paper is a problem of distributed power control for an interference channel, and for the multiple-access channel as a specific instance. A substantial line of works in the literature of distributed power control exploits game-theoretic tools to design power control schemes and to analyze their performance. One of such schemes is the iterative water-filling algorithm [16], which is a special instance of the best response dynamics (BRD), and is applied over a time horizon over which the channel state is constant. One of the main drawbacks of the various implementations of the BRD for power control problems, see e.g., [17], [18], [19], is that they tend to converge to Nash-equilibrium power control (NPC) policies. The latter are typically Pareto-inefficient, meaning that there exist some schemes which would allow all the agents to improve their individual utility with respect to (w.r.t.) the NPC policies. Another drawback is that such iterative schemes do not always converge. Only sufficient conditions for convergence are available, see e.g., [20] for the case of multiple input multiple output (MIMO) interference channels, and those conditions are sometimes very restrictive and even met with probability zero for some important special cases such as the parallel multiple-access channels [21]. In contrast, one of the main benefits of *coded power control*, which we develop in Section V, is precisely to obtain efficient operating points for the network. This is made possible by exchanging information about the channel state among the transmitters, and a key observation made in this paper is that this exchange can be achieved through observed quantities such as the signal-to-interference plus noise ratio (SINR). In other words, the SINRs of the different users can be viewed as the outputs of a channel over which transmitters communicate to coordinate their actions, provided they appropriately encode them. Coding appears because a transmitter maps several realizations of the channel state into a sequence of power levels, which then allows other transmitters to exploit the corresponding sequence of their SINRs to select their power levels. Since coding is used, no iterative procedure is required and convergence issues are therefore avoided. Since this paper focuses on efficiency, NPC will be compared to coded power control in terms of average sum-rate; other aspects such as CSI availability and complexity should also be considered but are deferred to future work.

B. Main contributions

The main contributions of the present work are as follows.

- First, the results reported in Section III extend [14] by relaxing some assumptions about the observation structure. Specifically, while [14] assumes that Agent 2 perfectly monitors the actions of Agent 1, we consider the case of imperfect monitoring. In addition we analyze both a situation similar to [14] in which Agent 2 has a strictly causal knowledge of the state (Theorem 8) and a situation in which Agent 2 knows nothing about the state (Theorem 14).
- Second, we clarify the connections between the game-theoretic formulation of [14] and information-theoretic considerations. In particular, links with coding theorems of the literature on state-dependent channels [22], [23], [24], [25], [26], separation theorems, and contributions on empirical coordination [4], [27] are discussed. In addition, the problem of determining the value of the long-run average payoff is formulated as an optimization problem, which is studied in detail in Section IV and exploited for the application to power control in Section V. This allows us to conduct a thorough numerical analysis for the problem of distributed power control and to assess the potential benefits of the proposed approach.
- Third, we establish a bridge between the problem of coordination via actions and the problem of power control in wireless networks. The present work develops a new perspective on resource allocation and control, in which designing a good resource allocation or control policy with high average common payoff amounts to designing a code. A good resource allocation or control code has to implement a trade-off between sending information about the upcoming realizations of the state, which play the role of messages and are required to obtain high payoff in the future, and achieving a good value for the current value of the payoff. As an illustration, Section V-E provides the complete description of a power control code for the multiple-access channel.

II. Problem statement

For the reader's convenience, we provide a non-exhaustive summary of the notation used throughout this paper in Table I.

We now formally introduce the problem studied in the remaining of the paper. We consider $K = 2$ agents that have to select their actions repeatedly over $T \geq 1$ stages or time-slots and wish to coordinate via their actions in the presence of a random state and with an observation structure detailed next. At each stage $t \in \{1, \cdots, T\}$, the action of Agent $k \in \{1, 2\}$ is $x_{k,t} \in \mathcal{X}_k$ with $|\mathcal{X}_k| < \infty$, while the realization of the random state is $x_{0,t} \in \mathcal{X}_0$ with $|\mathcal{X}_0| < \infty$. The realizations of the state are assumed to be i.i.d. according to a random variable X_0 with distribution ρ_0. The random state does not depend on the agents' actions but affects a common payoff function[3] $w : \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \to \mathbb{R}$. Coordination is measured in terms of the average payoff $W_T(x_0^T, x_1^T, x_2^T)$ as defined in (1). We assume that, at stage t, Agent 2 only has access to *imperfect* observations $y_t \in \mathcal{Y}$ of Agent 1's actions with $t \in \{1, \cdots, T\}$ and $|\mathcal{Y}| < \infty$. Specifically, the observations $y^T = \{y_1, \cdots, y_T\}$ are the output of a memoryless channel with transition probability

$$\mathrm{P}(y^T | x_0^T, x_1^T, x_2^T) = \prod_{t=1}^{T} \Gamma(y_t | x_{0,t}, x_{1,t}, x_{2,t}) \tag{3}$$

for some conditional probability Γ. We consider two asymmetric observation structures that restrict how agents observe the state and each other's actions. We characterize these structures through the strategies $(\sigma_t)_{1 \leq t \leq T}$ and $(\tau_t)_{1 \leq t \leq T}$ of Agents 1

[3]The function w can be any function such that the asymptotic average payoffs defined in the paper exist.

TABLE I
Summary of notation used throughout the paper.

Symbol	Meaning
Z	A generic random variable
Z_i^j	Sequence of random variables $(Z_i, ..., Z_j)$
$\mathcal{Z}$	Alphabet of Z
$\lvert\mathcal{Z}\rvert$	Cardinality of $\mathcal{Z}$
$\Delta(\mathcal{Z})$	Unit simplex over $\mathcal{Z}$
z	Realization of Z
z^n or $\underline{z}$	Sequence or vector $(z_1, z_2, \cdots, z_n)$
$\mathbb{E}_{\mathrm{P}}$	Expectation operator under the prob. P
$H(Z)$	Entropy of Z
$I(Y;Z)$	Mutual information between Y and Z
$Z_1 - Z_2 - Z_3$	Markov chain $\mathrm{P}(z_1 \lvert z_2, z_3) = \mathrm{P}(z_1 \lvert z_2)$
$\mathbb{1}_{\{\cdot\}}$	Indicator function
$\oplus$	Modulo-2 addition
$\mathbb{R}_+$	$[0, +\infty)$
$\mathcal{T}_\epsilon^n(Q)$	$\{x^n \in \mathcal{X}^n : \lVert T_{x^n} - Q \rVert_1 < \epsilon\}$

and 2, respectively, which are sequences of mappings defined for all stage $t \in \{1, \cdots, T\}$ as follows:

$$\text{case I:} \quad \left\{ \begin{array}{lcccc} \sigma_t^{\mathrm{I}} & : & \mathcal{X}_0^T & \to & \mathcal{X}_1 \\ \tau_t^{\mathrm{I}} & : & \mathcal{X}_0^{t-1} \times \mathcal{Y}^{t-1} & \to & \mathcal{X}_2 \end{array} \right. \tag{4}$$

$$\text{case II:} \quad \left\{ \begin{array}{lcccc} \sigma_t^{\mathrm{II}} & : & \mathcal{X}_0^T & \to & \mathcal{X}_1 \\ \tau_t^{\mathrm{II}} & : & \mathcal{Y}^{t-1} & \to & \mathcal{X}_2 \end{array} \right. . \tag{5}$$

Our objective is to characterize the set of average payoffs which are feasible asymptotically i.e., the possible values for $\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} w(x_{0,t}, x_{1,t}, x_{2,t})$ under the observation structures defined through (4) and (5). The definition of the two corresponding feasible sets is as follows.

Definition 1 (Feasible sets of payoffs)**.** *The feasible set of payoffs in case I is defined as*

$$\Omega^{\mathrm{I}} = \Big\{ \omega \in \mathbb{R} \ : \ \exists\, (\sigma_t^{\mathrm{I}}, \tau_t^{\mathrm{I}})_{1\le t\le T},\, \omega = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} w\left(x_{0,t}, \sigma_t^{\mathrm{I}}(x_0^T), \tau_t^{\mathrm{I}}(y^{t-1}, x_0^{t-1})\right) \Big\}. \tag{6}$$

The feasible set of payoffs in case II is defined as

$$\Omega^{\mathrm{II}} = \Big\{ \omega \in \mathbb{R} \ : \ \exists\, (\sigma_t^{\mathrm{II}}, \tau_t^{\mathrm{II}})_{1\le t\le T},\, \omega = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} w\left(x_{0,t}, \sigma_t^{\mathrm{II}}(x_0^T), \tau_t^{\mathrm{II}}(y^{t-1})\right) \Big\}. \tag{7}$$

The feasible sets of payoffs are directly related to the set of implementable empirical joint distributions over

$$\mathcal{X} \triangleq \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2, \tag{8}$$

defined as follows.

Definition 2 (Implementability)**.** *For $\ell \in \{\mathrm{I}, \mathrm{II}\}$, the probability distribution $\overline{Q}(x_0, x_1, x_2)$ is implementable if there exists a pair of strategies $(\sigma_t^\ell, \tau_t^\ell)_{1\le t\le T}$ inducing at each stage t a joint distribution*

$$\mathrm{P}_{X_{0,t}, X_{1,t}, X_{2,t}, Y_t}(x_0, x_1, x_2, y) \triangleq \Gamma(y|x_0, x_1, x_2) \mathrm{P}_{X_{1,t}, X_{2,t}|X_{0,t}}(x_1, x_2|x_0) \rho_0(x_0), \tag{9}$$

such that for all $(x_0, x_1, x_2) \in \mathcal{X}$

$$\lim_{T\to+\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{y\in\mathcal{Y}} \mathrm{P}_{X_{0,t}, X_{1,t}, X_{2,t}, Y_t}(x_0, x_1, x_2, y) = \overline{Q}(x_0, x_1, x_2). \tag{10}$$

Each feasible set of payoffs is the linear image of the corresponding set of implementable distributions under the expectation operator. Therefore, a certain value, say ω, is feasible (asymptotically) if and only if there exists an implementable distribution $\overline{Q}$ such that $\omega = \mathbb{E}_{\overline{Q}}[w] = \sum_{x_0, x_1, x_2} \overline{Q}(x_0, x_1, x_2) w(x_0, x_1, x_2)$. In the next section, we focus on the characterization of the implementable distributions rather than the direct characterization of the feasible sets payoffs.

The notion of implementable distribution may also be connected to the notion of achievable empirical distribution [4] upon introducing the *type* of the sequences of actions.

Definition 3 (Type [28])**.** *Let $T \ge 1$. For any sequence of realizations z^T of the generic random variable Z, the type of z^T, denoted by T_{z^T}, is the probability distribution on $\mathcal{Z}$ defined by*

$$\mathrm{T}_{z^T}(z) \stackrel{\triangle}{=} \frac{N(z|z^T)}{T} = \frac{1}{T} \sum_{t=1}^{T} \mathbb{1}_{\{z_t = z\}}, \tag{11}$$

where the notation $N(z|z^T)$ stands for the number of positions of z^T having the letter z.

By denoting T_{X^T} the distribution obtained by constructing the histogram of the sequence $x^T = (x_1, ..., x_T)$ with $x_t = (x_{0,t}, x_{1,t}, x_{2,t})$, it follows from Definition 2 that a distribution $\overline{Q}$ on $\mathcal{X}$ is implementable if there exists a sequence of strategies $(\sigma_t^\ell, \tau_t^\ell)_{1\le t\le T}$, $\ell \in \{\mathrm{I}, \mathrm{II}\}$ that generate, together with the sequence x_0^T, the sequence $x^T \in \mathcal{X}$ such that

$$\lim_{T\to\infty} ||\mathbb{E}(\mathrm{T}_{X^T}) - \overline{Q}||_1 = 0, \tag{12}$$

i.e., the average histogram of a sequence of actions is arbitrarily close to the distribution $\overline{Q}$.

Definition 4 (Empirical coordination [4])**.** *For $\ell \in \{\mathrm{I}, \mathrm{II}\}$, a distribution $\overline{Q}$ on $\mathcal{X}$ is an achievable empirical coordination if there exists a sequence of strategies $(\sigma_t^\ell, \tau_t^\ell)_{1\le t\le T}$ that generate, together with the sequence x_0^T, the sequence $x^T \in \mathcal{X}$ such that*

$$\forall \epsilon > 0, \quad \lim_{T\to\infty} P(||\mathrm{T}_{X^T} - \overline{Q}||_1 > \epsilon) = 0, \tag{13}$$

i.e., the distance between the histogram of a sequence of actions and $\overline{Q}$ converges in probability to 0.

Proposition 5. *If $\overline{Q} \in \Delta(\mathcal{X})$ is an achievable empirical coordination, then it is implementable.*

The proof of Proposition 5 is provided in Appendix A.

Before analyzing the considered coordination problem in details in Section III, a few important comments are in order. In [14], the authors assume that the strategies are given by the two sequences of mappings $\sigma_t : \mathcal{X}_0^T \times \mathcal{X}_1^{t-1} \times \mathcal{X}_2^{t-1} \to \mathcal{X}_1$ and $\tau_t : \mathcal{X}_0^{t-1} \times \mathcal{X}_1^{t-1} \times \mathcal{X}_2^{t-1} \to \mathcal{X}_2$. In particular, this means that Agent 2 perfectly monitors the actions of Agent 1 i.e., $Y = X_1$. One of our results is that, under perfect monitoring, there is no loss in terms of feasibility by considering strategies given by (5) instead of those assumed in [14]. However, the set of possible Nash equilibrium points obtained with (5) may not coincide with the one associated with the strategies of [14]. Characterizing the set of possible Nash equilibrium points is relevant when agents have diverging interests. It turns out that when this is the case, the fact that an agent can observe the actions of the other agents matters. In a power control setting for instance, if the transmitters implement a cooperation plan that consists in transmitting at low power as long as no transmitter uses a high power level, see e.g., [29], it matters for the transmitters to be able to check whether the other transmitters effectively use a low power level.

Note that the definition of the strategies are not block strategies, as assumed for conventional coding and decoding. Here, an agent acts at every stage, and the considered strategies can be seen as joint source-channel coding strategies with an online coding and decoding requirement.

Finally, we emphasize that the strategy for Agent 1 assumes a non-causal knowledge of the state. Such assumptions have been largely used in the information theory literature, going back to the work of Gel'fand and Pinsker [22] on coding over state-dependent channels. For instance, Gel'fand-Pinsker coding is fully relevant for the problem of coding over broadcast channels [30][31] and for watermarking problems [32]. Since one of the main contributions of this paper detailed in Section V is the application of the developed framework to power control, we provide here a few additional practical motivations for such an assumption. First, even if Agent 1 only knows the future realizations for a reduced time horizon[4], the network can be coordinated to a degree which offers a significant performance gain compared to conventional approaches, such as implementing Nash equilibrium-type distributed policies, see e.g., [33], [16], [20], [17]. The most typical situation in power control is to assume that two phases are available, a training phase and an action phase, and that one state is known in advance to adjust the power level. This special case corresponds to setting $T = 2$ that is, $t \in \{1, 2\}$. In such a scenario, and as illustrated in Fig. 1, a simple coordination strategy when Agent 1 knows the upcoming channel state might be to inform Agent 2 about it[5] on odd time-slots and coordinate their actions on even time-slots. Note that assuming that Agent 1 knows all the realizations of the state in advance can also be seen as a way of obtaining an upper bound for the performance of scenarios that assume a reduced time horizon for forecasting. It is worth noting that practical scenarios have appeared over the recent years, for which forecasting a state over a long time horizon is realistic. For instance, it has become more and more common to exploit the forecast trajectory of a mobile user to optimize the system [34], which also makes our approach relevant when the channel state is interpreted as the path loss. Finally, the proposed approach may also be applied to the case where the state is not i.i.d. from stage to stage, but is i.i.d. from block to block, where a block consists of several stages. Indeed, there exist wireless communication standards that assume the channel to be constant over several time-slots, and the proposed approach suggests that gains can be obtained by varying the power level from time-slot to time-slot even if the channel is constant.

III. Information constraints on implementable empirical joint distributions

In this section, we characterize the sets of implementable empirical joint distributions $\overline{Q} \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ for the cases described through (4) and (5). We show that these sets consist of distributions in $\Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ subject to an *information constraint*, which captures the restrictions imposed on the observation structure between the agents. Specifically, Theorem 6 in Section III-A provides a necessary condition for implementability, while Theorem 8 and Theorem 14 in Section III-B provide sufficient conditions for implementability in case I and case II, respectively.

A. A necessary condition for implementability

Theorem 6. *Let $\overline{Q}$ be a distribution in $\Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ such that $\forall\, x_0 \in \mathcal{X}_0$, $\sum_{x_1,x_2} \overline{Q}(x_0, x_1, x_2) = \rho_0(x_0)$. In both case I and case II, a distribution $\overline{Q}$ is implementable if it is the marginal of a distribution $Q \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})$ factorizing as*

$$Q(x_0, x_1, x_2, y) = \Gamma(y|x_0, x_1, x_2)\overline{Q}(x_0, x_1, x_2),\ (x_0, x_1, x_2, y) \in \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y} \tag{14}$$

and satisfying the information constraint

$$I_Q(X_0; X_2) \leq I_Q(X_1; Y|X_0, X_2) \tag{15}$$

[4]In a power control problem, knowing only one realization ahead is already very useful.

[5]For example, Transmitter 1 might use a high (resp. low) power level on an odd time-slot to inform Transmitter 2 that the channel is good (resp. bad) in the next even time-slot.

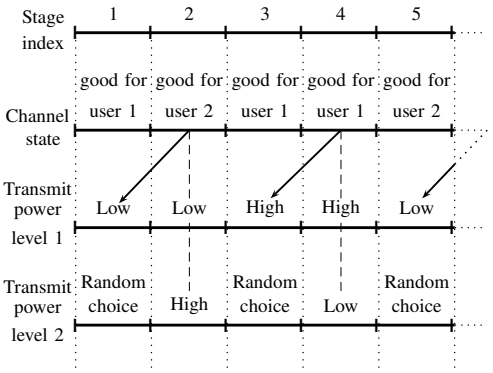

Fig. 1. This figure illustrates a simple coordination scheme between two transmitters (which are the agents) in a simplified scenario where the alphabets are binary: $\mathcal{X}_0 = \{\text{good for user 1}, \text{good for user 2}\}$, $\mathcal{X}_1 = \{\text{low}, \text{high}\}$, $\mathcal{X}_2 = \{\text{low}, \text{high}\}$. The informed transmitter (i.e., 1) chooses the lowest (resp. highest) transmit power on the current stage $2t'+1$ if the upcoming channel state on stage $2t'+2$ is good for user 2 (resp. 1). If Transmitter 2 can perfectly retrieve the power levels of Transmitter 1, it therefore knows the realization of the state on stages whose index is even. It transmits at low (resp. high) power if the channel is good for user 1 (resp. 2). For stages whose index is odd, it chooses its power at random.

where

$$I_Q(X_0; X_2) = \sum_{x_0, x_2} Q_{X_0X_2}(x_0, x_2) \cdot \log \frac{Q_{X_0X_2}(x_0, x_2)}{\rho_0(x_0).\sum_{x_0} Q_{X_0X_2}(x_0, x_2)}, \tag{16}$$

$$Q_{X_0X_2}(x_0, x_2) = \sum_{x_1} \overline{Q}(x_0, x_1, x_2), \tag{17}$$

$$I_Q(X_1; Y|X_0, X_2) = \sum_{x_0, x_1, x_2, y} Q(x_0, x_1, x_2, y) \times \left[\log \frac{Q(x_0, x_1, x_2, y).Q_{X_0X_2}(x_0, x_2)}{\sum_y Q(x_0, x_1, x_2, y).\sum_{x_1} Q(x_0, x_1, x_2, y)} \right]. \tag{18}$$

Proof: Let $\overline{Q} \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ be an implementable distribution according to Definition 2. Then for $\ell \in \{\text{I}, \text{II}\}$, there exists $(\sigma_t^\ell, \tau_t^\ell)_{1 \le t \le T}$ such that for all (x_0, x_1, x_2)

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{y \in \mathcal{Y}} P_{X_{0,t}, X_{1,t}, X_{2,t}, Y_t}(x_0, x_1, x_2, y) = \overline{Q}(x_0, x_1, x_2), \tag{19}$$

where $P_{X_{0,t}, X_{1,t}, X_{2,t}, Y_t}$ is defined in (9). Because of the specific form of (9), this also implies that

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} P_{X_{0,t}, X_{1,t}, X_{2,t}, Y_t}(x_0, x_1, x_2, y) = Q(x_0, x_1, x_2, y) \tag{20}$$

with Q as in (14). We now establish the information constraint in (15) that must be satisfied by Q through a series of information-theoretic inequalities. We have

$$\sum_{t=1}^{T} I_{P_{X_{0,t}, X_{1,t}, X_{2,t}, Y_t}}(X_0; X_2) = \sum_{t=1}^{T} I(X_{0,t}; X_{2,t}) \tag{21}$$

$$= \sum_{t=1}^{T} H(X_{0,t}) - H(X_{0,t}|X_{2,t}) \tag{22}$$

$$\overset{(a)}{=} H(X_0^T) - \sum_{t=1}^{T} H(X_{0,t}|X_{2,t}) \tag{23}$$

$$= H(X_0^T, Y^T, X_2^T) - H(Y^T, X_2^T|X_0^T) - \sum_{t=1}^{T} H(X_{0,t}|X_{2,t}) \tag{24}$$

$$= H(X_0^T, Y^T, X_2^T) - H(X_2^T|X_0^T) - H(Y^T|X_0^T, X_2^T) - \sum_{t=1}^{T} H(X_{0,t}|X_{2,t}) \tag{25}$$

$$= H(X_0^T, Y^T, X_2^T) - H(X_2^T|X_0^T) - H(Y^T|X_0^T, X_2^T) + H(Y^T|X_0^T, X_1^T, X_2^T) - \sum_{t=1}^{T} H(X_{0,t}|X_{2,t}) \tag{26}$$

$$\overset{(b)}{=} \sum_{t=1}^{T} \Big[H(X_{0,t}, Y_t, X_{2,t}|X_0^{t-1}, Y^{t-1}, X_2^{t-1}) - H(X_{0,t}|X_{2,t}) - H(Y_t|X_{0,t}, X_{1,t}, X_{2,t}) \Big] + H(Y^T|X_0^T, X_1^T, X_2^T) - H(X_2^T|X_0^T) - H(Y^T|X_0^T, X_2^T) \tag{27}$$

$$= \sum_{t=1}^{T} \Big[H(X_{0,t}, Y_t, X_{2,t}|X_0^{t-1}, Y^{t-1}, X_2^{t-1}) - H(X_{0,t}|X_{2,t}) - H(Y_t|X_{0,t}, X_{1,t}, X_{2,t}) + H(Y_t|X_{0,t}, X_{2,t}) - H(Y_t|X_{0,t}, X_{2,t}) \Big] + H(Y^T|X_0^T, X_1^T, X_2^T) - H(X_2^T|X_0^T) - H(Y^T|X_0^T, X_2^T) \tag{28}$$

$$= \sum_{t=1}^{T} \Big[H(Y_t|X_{0,t}, X_{2,t}) - H(Y_t|X_{0,t}, X_{1,t}, X_{2,t}) + H(X_{0,t}, Y_t, X_{2,t}|X_0^{t-1}, Y^{t-1}, X_2^{t-1}) - H(X_{0,t}|X_{2,t}) - H(Y_t|X_{0,t}, X_{2,t}) \Big] + H(Y^T|X_0^T, X_1^T, X_2^T) - H(X_2^T|X_0^T) - H(Y^T|X_0^T, X_2^T) \tag{29}$$

$$= \sum_{t=1}^{T} \Big[\underbrace{H(Y_t|X_{0,t}, X_{2,t}) - H(Y_t|X_{0,t}, X_{1,t}, X_{2,t})}_{A_t} + \underbrace{H(X_{0,t}|X_0^{t-1}, Y^{t-1}, X_2^{t-1}) - H(X_{0,t}|X_{2,t})}_{B_t} + \underbrace{H(Y_t|X_{0,t}, X_{2,t}, X_0^{t-1}, Y^{t-1}, X_2^{t-1}) - H(Y_t|X_{0,t}, X_{2,t})}_{C_t} + \underbrace{H(X_{2,t}|X_{0,t}, X_0^{t-1}, Y^{t-1}, X_2^{t-1})}_{D_t} \Big] + \underbrace{H(Y^T|X_0^T, X_1^T, X_2^T) - H(Y^T|X_0^T, X_2^T) - H(X_2^T|X_0^T)}_{E_T}, \tag{30}$$

where (a) follows because $(X_{0,t})_{1\leq t\leq T}$ is i.i.d. and (b) from the fact that the observations Y^T are the output of a memoryless channel per (3). Note that, for all t, $C_t \leq 0$ and $E_t \leq 0$ since conditioning reduces entropy; $B_t \leq 0$ whether we define $X_{2,t} = \tau_t^{\mathrm{I}}(X_0^{t-1}, Y^{t-1})$ or $X_{2,t} = \tau_t^{\mathrm{II}}(Y^{t-1})$; $D_t = 0$ whether we define $X_{2,t} = \tau_t^{\mathrm{I}}(X_0^{t-1}, Y^{t-1})$ or $X_{2,t} = \tau_t^{\mathrm{II}}(Y^{t-1})$. Therefore,

$$\sum_{t=1}^{T} I_{P_{X_{0,t}, X_{1,t}, X_{2,t}, Y_t}}(X_0; X_2) \leq \sum_{t=1}^{T} H(Y_t|X_{0,t}, X_{2,t}) - H(Y_t|X_{0,t}, X_{1,t}, X_{2,t}) \tag{31}$$

$$= \sum_{t=1}^{T} I(X_{1,t}; Y_t|X_{0,t}, X_{2,t}) \tag{32}$$

$$= \sum_{t=1}^{T} I_{P_{X_{0,t}, X_{1,t}, X_{2,t}, Y_t}}(X_1; Y|X_0, X_2). \tag{33}$$

Let us introduce the function Φ^{I}:

$$\begin{array}{rccc} \Phi^{\mathrm{I}}: & \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}) & \to & \mathbb{R} \\ & Q & \mapsto & I_Q(X_0; X_2) - I_Q(X_1; Y|X_0, X_2) \end{array}. \tag{34}$$

It follows from (31)-(33) that

$$\sum_{t=1}^{T} \Phi^{\mathrm{I}}(P_{X_{0,t}, X_{1,t}, X_{2,t}, Y_t}) \leq 0. \tag{35}$$

To conclude the proof, we exploit the following lemma, which proof can be found in Appendix B.

Lemma 7. *The function Φ^{I} is strictly convex over the set of distributions $Q \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})$ that verify*

$$\forall x_0 \in \mathcal{X}_0, \quad \sum_{x_1,x_2,y} Q(x_0,x_1,x_2,y) = \rho_0(x_0), \tag{36}$$

and $\forall\,(x_0,x_1,x_2,y) \in (\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})$:

$$Q(x_0,x_1,x_2,y) = \Gamma(y|x_0,x_1,x_2)\rho_0(x_0)Q(x_1,x_2|x_0), \tag{37}$$

with ρ_0 and Γ fixed.

Let us define the joint distribution $P^{(T)}_{X_0,X_1,X_2,Y}$ as

$$P^{(T)}_{X_0,X_1,X_2,Y}(x_0,x_1,x_2,y) \triangleq \frac{1}{T}\sum_{t=1}^{T}\sum_{y\in\mathcal{Y}} P_{X_{0,t},X_{1,t},X_{2,t},Y_t}(x_0,x_1,x_2,y). \tag{38}$$

Since Φ^{I} is convex by Lemma 7, we know that

$$\Phi^{\mathrm{I}}(P^{(T)}_{X_0,X_1,X_2,Y}) \leq \frac{1}{T}\sum_{t=1}^{T}\Phi^{\mathrm{I}}(P_{X_{0,t},X_{1,t},X_{2,t},Y_t}). \tag{39}$$

In addition, since Φ^{I} is continuous and because of (20), $\forall\varepsilon' > 0$, there exists T' such that $\forall T \geq T'$,

$$\Phi^{\mathrm{I}}(Q) \leq \Phi^{\mathrm{I}}(P^{(T)}_{X_0,X_1,X_2,Y}) + \varepsilon'. \tag{40}$$

Therefore, combining (35), (38), and (40), we obtain

$$\forall\varepsilon' > 0, \quad \Phi^{\mathrm{I}}(Q) \leq 0, \tag{41}$$

which concludes the proof. ∎

Theorem 6 can be interpreted as follows. Agent 2's actions, which are represented by X_2, correspond to a joint source-channel decoding operation with distortion on the information source, which is represented by X_0. To be achievable, the distortion rate has to be less than the transmission rate allowed by the channel, whose input and output are represented by Agent 1's action X_1 and the signal Y observed by Agent 2. Therefore, the pair $S = (X_0, X_2)$ seems to play the same role as the side information in state-dependent channels [31]. Although we exploit this interpretation when establishing sufficient conditions for implementability in Section III-B, the argument seems inappropriate to show that the sufficient conditions are also necessary. In contrast to classical arguments in converse proofs for state-dependent channels [22], [35], in which the transmitted "message" is independent of the state, here the role of the message is played by the quantity X_0^T, which is not independent of $S^T = (X_0^T, X_2^T)$. This is one of the reasons why the converse proof does not follow from existing results.

B. Sufficient conditions for implementability

We start by characterizing implementable distributions for the observation structure corresponding to case I in (4).

Theorem 8. *Consider the observation structure in case I. Let $\overline{Q} \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ be such that $\sum_{(x_1,x_2)} \overline{Q}(x_0,x_1,x_2) = \rho_0(x_0)$ for all $x_0 \in \mathcal{X}_0$. If the distribution Q defined as*

$$Q(x_0,x_1,x_2,y) = \Gamma(y|x_0,x_1,x_2)\overline{Q}(x_0,x_1,x_2) \tag{42}$$

for $(x_0,x_1,x_2,y) \in \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}$ verifies the constraint

$$I_Q(X_0;X_2) \leq I_Q(X_1;Y|X_0,X_2), \tag{43}$$

then $\overline{Q}$ is implementable in the sense of Definition 2.

Proof: Consider a joint distribution $Q_{X_0X_1X_2Y} \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})$ that factorizes as in (42). The crux of the proof is to design the strategies $f : \underline{x}_0^{(b+1)}, \underline{x}_0^{(b)}, \underline{x}_2^{(b)} \mapsto \underline{x}_1^{(b)}$ and $g : \underline{x}_0^{(b)}, \underline{y}^{(b)}, \underline{x}_2^{(b)} \mapsto \underline{x}_2^{(b+1)}$ from a block-Markov coding scheme that operates over B blocks of $n \geq 1$ actions. As illustrated in Fig. 2, in every block $1 \leq b \leq B-1$, Agent 1 will communicate to Agent 2 the actions that Agent 2 should play in the next block $b+1$. This is made possible by restricting the sequence of actions played by Agent 2 in each block b to a codebook of actions $(\underline{x}_2^{(b)}(i_b))_{1\leq i_b\leq 2^{nR}}$ with $i_b \in \{1,\ldots,2^{nR}\}$, so that Agent 1 only has to communicate the index i_b to be played in the next block. The problem then reduces to a source-channel coding problem over a state dependent channel, for which in every block b:

- Agent 1 communicates with Agent 2 over a state-dependent discrete memoryless channel (DMC) with transition probability $\Gamma(y|x_0,x_1,x_2)$;
- the state of the channel consists of the state sequence $\underline{x}_0^{(b)}$ and the current action sequence $\underline{x}_2^{(b)}(i_{b-1})$ for Agent 2;

- the state is known non-causally by Agent 1 and causally by Agent 2, to be consistent with the observation structure defined in (4);
- the source sequence to be communicated is a sequence $\underline{x}_2^{(b+1)}(i_b)$, chosen to be empirically coordinated with the future state sequence $\underline{x}_0^{(b+1)}$ in block $b+1$;
- i_b is encoded in a codeword $\underline{x}_1^{(b)}$ chosen to empirically coordinated with the current state $(\underline{x}_0^{(b)}, \underline{x}_2^{(b)}(i_{b-1}))$.

Intuitively, R must be sufficiently large so that one may find a codeword $\underline{x}_2^{(b)}(i_b)$ coordinated with any state sequence $\underline{x}_0^{(b)}$; simultaneously, R must be small enough to ensure that the codeword $\underline{x}_2^{(b)}(i_b)$ can be reliably decoded by Agent 2 after transmission over the channel $\Gamma(y|x_0, x_1, x_2)$. The formal analysis of these conditions, which we develop next, will establish the result.

Unlike the block-Markov schemes used, for instance, in relay channels, in which all nodes may agree on a *fixed* message in the first block at the expense of a small rate loss, the first block must be dealt with more carefully in the case of coordination. In fact, Agent 1 must transmit a message defined by the state realization $\underline{x}_0^{(2)}$, which is unknown to Agent 2. Consequently, we have to account for an "uncoordinated" transmission in the first block, in which Agent 1 is forced to communicate at rate $\widehat{R}$ that differs from the rate R used in subsequent blocks. To characterize $\widehat{R}$, we introduce another joint distribution $\widehat{Q} \in \Delta(\mathcal{X} \times \mathcal{Y})$ that factorizes as

$$\widehat{Q}(x_0, x_1, x_2, y) = \Gamma(y|x_0, x_1, x_2)\overline{Q}(x_1|x_0, x_2)\widehat{Q}(x_0)\widehat{Q}(x_2). \tag{44}$$

Note that $\widehat{Q}$ differs from Q in that X_0 and X_2 are independent.

Let $\epsilon > 0$. Let $R, \widehat{R} > 0$, $\frac{\epsilon}{2} > \epsilon_2 > \epsilon_1 > 0$ be real numbers and $n \in \mathbb{N}^*$ to be specified later, and define

$$\alpha \triangleq \max\left(\lceil R/\widehat{R} \rceil, 1\right) \tag{45}$$

$$B \triangleq \left\lceil 1 + \alpha\left(\frac{4}{\epsilon} - 1\right)\right\rceil. \tag{46}$$

Intuitively, α measures the rate penalty suffered from the uncoordinated transmission at rate $\widehat{R}$ in the first block. The choice of B merely ensures that $\frac{2\alpha}{B-1+\alpha} \leq \frac{\epsilon}{2}$, as exploited later.

Source codebook generation for $b = 1$. Choose a sequence $\underline{x}_2^{(1)}(i_0) \in \mathcal{T}_{\epsilon_1}^n(Q_{X_2})$. This unique sequence is revealed to both agents.

Source codebooks generation for $b \in \{2, \ldots, B+1\}$. Randomly and independently generate $\lceil 2^{nR} \rceil$ sequences according to $\Pi_{t=1}^n Q_{X_2}(x_{2,t})$. Label the sequences $\underline{x}_2^{(b)}(i)$ with $i \in \{1, \ldots, 2^{nR}\}$. These sequences constitute the source codebook $\mathcal{C}_s^{(b)}$, which is revealed to both agents.

Channel codebook generation for $b = 1$. Following [31, Section 7.4.1], generate a codebook $\mathcal{C}_c^{(1)}$ of $\lceil 2^{\alpha n \widehat{R}} \rceil$ codewords and of length αn for a channel with random state (x_0, x_2) available causally at both Agent 1 and Agent 2. With a slight abuse of notation, denote the codewords corresponding to a message i and a state sequence $(\underline{x}_0, \underline{x}_2^{(1)}(i_0))$ as $\underline{x}_1^{(b)}(i, \underline{x}_0, i_0)$.

Channel codebooks generation for $b \in \{2, \ldots, B\}$. For every $(\underline{x}_0, j) \in \mathcal{X}_0^n \times \{1, \ldots, 2^{nR}\}$, randomly and independently generate $\lceil 2^{nR} \rceil$ sequences according to $\Pi_{t=1}^n Q_{X_1|X_0,X_2}(x_{1,t}|x_{0,t}, x_{2,t}^{(b)}(j))$, where $x_{2,t}^{(b)}(j)$ is the t-th component of $\underline{x}_2^{(b)}(j) \in \mathcal{C}_s^{(b)}$. Label the sequences $\underline{x}_1^{(b)}(i, \underline{x}_0, j)$ with $i \in \{1, \ldots, 2^{nR}\}$, $\underline{x}_0 \in \mathcal{X}_0^n$, and $j \in \{1, \ldots, 2^{nR}\}$. These sequences constitute the channel codebook $\mathcal{C}_c^{(b)}$, which is revealed to both agents.

Source and channel encoding at Agent 1 in block $b \in \{1, \ldots, B\}$. At the beginning of block b, Agent 1 uses his non-causal knowledge of the state $\underline{x}_0^{(b+1)}$ in the next block $b+1$ to looks for an index i_b such that $(\underline{x}_0^{(b+1)}, \underline{x}_2^{(b+1)}(i_b)) \in \mathcal{T}_{\epsilon_1}^n(Q_{X_0X_2})$. If there is more than one such index, it chooses the smallest among them, otherwise it chooses $i_b = 1$. Agent 1 then encodes i_b

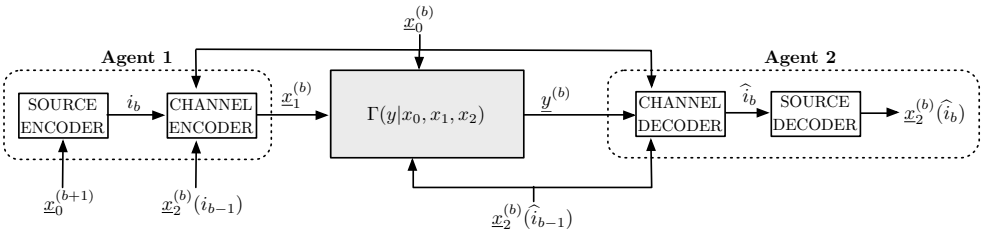

Fig. 2. Illustration of encoding and decoding operations in block b of the proposed block-coding scheme.

using its knowledge of the channel state sequence $(\underline{x}_0^{(b)}, \underline{x}_2^{(b)}(i_{b-1}))$ as $\underline{x}_1^{(b)}(i_b, \underline{x}_0^{(b)}, i_{b-1})$. Note that this is possible in block 1 because of the specific codebook and of the choice of α in (45).

Channel decoding at Agent 2 in block $b = 1$. At the end of the block, Agent 2 uses the same decoding procedure as in [31, Section 7.4.1]; it uses the channel output $\underline{y}^{(1)}$ and its knowledge of the channel state sequence $(\underline{x}_0^{(1)}, \underline{x}_2^{(1)}(i_0))$ to form an estimate $\widehat{i}_1$ of i_1.

Channel decoding at Agent 2 in block $b \in \{2, \dots, B\}$. At the end of block b, Agent 2 knows the sequence of channel outputs $\underline{y}^{(b)}$ and the channel state sequence $(\underline{x}_0^{(b)}, \underline{x}_2^{(b)}(\widehat{i}_{b-1}))$ in block b. Agent 2 then looks for an index $\widehat{i}_b$ such that

$$(\underline{x}_0^{(b)}, \underline{x}_1^{(b)}(\widehat{i}_b, \underline{x}_0^{(b)}, \widehat{i}_{b-1}), \underline{y}^{(b)}, \underline{x}_2^{(b)}(\widehat{i}_{b-1})) \in \mathcal{T}_{\epsilon_2}^n(Q_{X_0X_1X_2Y}). \tag{47}$$

If there is none of more than one such index, Agent 2 sets $\widehat{i}_b = 1$.

Source decoding at Agent 2 in block $b \in \{1, \dots, B\}$. Agent 2 transmits $\underline{x}_2^{(b)}(\widehat{i}_{b-1})$, where $\widehat{i}_{b-1}$ is its estimate of the message transmitted by Agent 1 in the previous block $b-1$, with the convention that $\widehat{i}_0 = i_0$.

Analysis. According to Proposition 5, it is sufficient to prove that $\overline{Q}$ is an achievable empirical coordination. We therefore introduce the event

$$E \triangleq \{(\underline{X}_0^T, \underline{X}_1^T, \underline{X}_2^T) \notin \mathcal{T}_\epsilon^n(\overline{Q})\} \tag{48}$$

and we proceed to show that $P(E)$ can be made arbitrarily small for n sufficiently large and a proper choice of the rates R and $\widehat{R}$. We start by introducing the following events in each block $b \in \{1, \dots, B\}$.

$$\begin{aligned}
E_0 &\triangleq \{I_1 \neq \widehat{I}_1\} \\
E_1^{(b)} &\triangleq \{(\underline{X}_0^{(b+1)}, \underline{x}_2^{(b)}(i)) \notin \mathcal{T}_{\epsilon_1}^n(Q_{X_0X_2}) \, \forall \, i \in \{1, 2, \dots, 2^{nR}\}\} \\
E_2^{(b)} &\triangleq \{(\underline{X}_0^{(b)}, \underline{x}_1^{(b)}(I_b, \underline{X}_0^{(b)}, \widehat{I}_{b-1}), \underline{x}_2^{(b)}(\widehat{I}_{b-1}), \underline{Y}^{(b)}) \notin \mathcal{T}_{\epsilon_2}^n(Q)\} \\
E_3^{(b)} &\triangleq \left\{(\underline{X}_0^{(b)}, \underline{x}_1^{(b)}(i, \underline{X}_0^{(b)}, \widehat{I}_{b-1}), \underline{x}_2^{(b)}(\widehat{I}_{b-1}), \underline{Y}^{(b)}) \in \mathcal{T}_{\epsilon_2}^n(Q) \text{ for some } i \neq I_b\right\}.
\end{aligned}$$

We start by developing an upper bound for $\|T_{\underline{x}_0^T \underline{x}_1^T \underline{x}_2^T} - \overline{Q}\|_1$, whose proof can be found in Appendix C.

Lemma 9.

$$\|T_{\underline{x}_0^T \underline{x}_1^T \underline{x}_2^T} - \overline{Q}\|_1 \leq \frac{2\alpha}{B-1+\alpha} + \frac{1}{B-1} \sum_{b=2}^{B} \|T_{\underline{x}_0^{(b)} \underline{x}_1^{(b)} \underline{x}_2^{(b)}} - \overline{Q}\|_1, \tag{49}$$

Recalling the choice of B in (46), we therefore have

$$\begin{aligned}
\mathrm{P}(E) &= \mathrm{P}\left(\|T_{\underline{X}_0^T \underline{X}_1^T \underline{X}_2^T} - \overline{Q}\|_1 \geq \epsilon\right) && (50) \\
&\leq \mathrm{P}\left(\frac{1}{B-1} \sum_{b=2}^{B} \|T_{\underline{X}_0^{(b)} \underline{X}_1^{(b)} \underline{X}_2^{(b)}} - \overline{Q}\|_1 \geq \frac{\epsilon}{2}\right) && (51) \\
&\leq \mathrm{P}\left(\|T_{\underline{X}_0^{(b)} \underline{X}_1^{(b)} \underline{X}_2^{(b)}} - \overline{Q}\|_1 \geq \frac{\epsilon}{2} \text{ for some } b \in \{2, \dots, B\}\right) && (52) \\
&\leq \mathrm{P}\left(E_0 \cup E_1^{(1)} \bigcup_{b=2}^{B} \left(E_1^{(b)} \cup E_2^{(b)} \cup E_3^{(b)}\right)\right) && (53) \\
&\leq \mathrm{P}\left(E_0\right) + \sum_{b=1}^{B} \mathrm{P}\left(E_1^{(b)}\right) + \sum_{b=2}^{B} \left(\mathrm{P}\left(E_2^{(b)} | E_1^{(b-1)c} \cap E_2^{(b-1)c} \cap E_3^{(b-1)c} \cap E_0^c\right)\right) \\
&\qquad + \sum_{b=2}^{B} \mathrm{P}\left(E_3^{(b)} | E_1^{(b-1)c} \cap E_2^{(b-1)c} \cap E_3^{(b-1)c} \cap E_0^c\right) && (54)
\end{aligned}$$

As proved in Appendix C, the following lemmas show that all the averages over the random codebooks of the term above vanish as $n \to \infty$.

Lemma 10. *If $\widehat{R} < (1-\epsilon_2) I_{\widehat{Q}}(X_1; Y | X_0, X_2) - \delta(\epsilon_2)$, then*

$$\lim_{n \to \infty} \mathbb{E}\left(\mathrm{P}(E_0)\right) = 0. \tag{55}$$

Lemma 11. *If $R > I_Q(X_0; X_2) + \delta(\epsilon_1)$, then for any $b \in \{1, \dots, B\}$,*

$$\lim_{n \to \infty} \mathbb{E}\left(\mathrm{P}(E_1^{(b)})\right) = 0. \tag{56}$$

Lemma 12. *For any $b \in \{2, \dots, B\}$,*

$$\lim_{n\to\infty} \mathbb{E}\left(\mathrm{P}(E_2^{(b)}|E_1^{(b-1)c} \cap E_2^{(b-1)c} \cap E_3^{(b-1)c} \cap E_0^c)\right) = 0. \tag{57}$$

Lemma 13. *If $R < I_Q(X_1; Y|X_0, X_2) - \delta(\epsilon_2)$, then for any $b \in \{2, \dots, B\}$,*

$$\lim_{n\to\infty} \mathbb{E}\left(\mathrm{P}(E_3^{(b)})\right) = 0. \tag{58}$$

Therefore, if $I_{\widehat{Q}}(X_1; Y|X_0, X_2) > 0$ and $I_Q(X_0; X_2) < I_Q(X_1; Y|X_0, X_2)$, we can find ϵ_1, ϵ_2 small enough such that $R > 0$, $I_Q(X_0; X_2) < R < I_Q(X_1; Y|X_0, X_2)$, and

$$\lim_{n\to\infty} \mathbb{E}\left(\mathrm{P}(E)\right) = 0. \tag{59}$$

Hence, there must exist at least one set of codes such that $\lim_{n\to\infty} \mathrm{P}(E) = 0$. Since $\epsilon > 0$ can be chosen arbitrarily small, $\overline{Q}$ is an achievable empirical coordination. ■

A few comments are in order regarding the coding scheme achieving the result in Theorem 8. Note that Agent 1 uses $(\underline{w}_0^{(b)}, \underline{w}_2^{(b)}(i_{b-1}))$ as "side information" to correlate its actions during block b, while Agent 2 uses $(\underline{w}_0^{(b)}, \underline{w}_2^{(b)}(\widehat{i}_{b-1}))$. However, under the condition $I_Q(X_0; X_2) < I_Q(X_1; Y|X_0, X_2)$, $\widehat{i}_{b-1} = i_{b-1}$ with overwhelming probability as the block length n goes to infinity. Hence, Agent 1 effectively knows the actions of Agent 2 even though it does not directly observe them. Furthermore, note that the state sequence $\underline{w}_0^{(b+1)}$, which plays the role of the message to be encoded in block b, is independent of the "side information" $(\underline{w}_0^{(b)}, \underline{w}_2^{(b)}(i_{b-1}))$; this allows us to readily reuse classical coding schemes for the transmission of messages over state-dependent channels.

Apart for the strict inequality in (43) instead of the inequality in (15), the information constraints of Theorem 6 and Theorem 8 coincide, hence establishing a necessary and sufficient condition for a joint distribution $\overline{Q}(x_0, x_1, x_2)$ to be implementable. This also shows that having Agent 1 select the actions played by Agent 2 and separating source and channel encoding operations do not incur any loss of optimality. From Theorem 6 and Theorem 8 one can therefore provide the complete characterization of the set of achievable payoffs for the observation structure defined in (4) and for an i.i.d. random state. As further explored in Section V, one can apply this result to an interference network with two transmitters and two receivers, in which Transmitter 1 may represent the most informed agent, such as a primary transmitter [36], [37], Γ may represent an SINR feedback channel from Receiver 2 to Transmitter 2.

While Theorem 8 has been derived for the case of an i.i.d. random state, the result generalizes to a situation in which the state is constant over $L \geq 1$ consecutive stages and i.i.d. from one stage to the next. In such case, the information constraint becomes

$$\frac{1}{L} I_Q(X_0; X_2) < I_Q(X_1; Y|X_0, X_2) \tag{60}$$

This modified constraint is useful in some setting, such as in wireless communications for which channels can often be assumed to be block i.i.d.. When $L \to \infty$, which correspond to a single realization of the random state, the information constraint is always satisfied and any joint distribution over $\Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ is implementable.

Finally, we emphasize that the information constraint obtained when coordinating via actions differs from what would be obtained when coordinating using classical communication [38] with a dedicated channel. If Agent 1 could communicate with Agent 2 through a channel with capacity C, then all joint distributions $\widetilde{Q} \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ subject to the information constraint

$$I_{\widetilde{Q}}(X_0; X_2) \leq C \tag{61}$$

would be implementable.

In contrast, the constraint $I_Q(X_0; X_2) < I_Q(X_1; Y|X_0X_2)$ reflects the following two distinctive characteristics of communication via actions.

1) The input distribution X_1 to the "implicit channel" used for communication between Agent 1 and Agent 2 cannot be optimized independently of the actions and of the global state.
2) The output Y of the implicit channel depends not only on X_1 but also on (X_0, X_2); essentially, the global state X_0 and the actions X_2 of Agent 2 act as a state for the implicit channel.

Under specific conditions, the coordination via actions may reduce to coordination with a dedicated channel. For instance, if the payoff function factorizes as $w(x_0, x_1, x_2) \triangleq w_1(x_1)w_2(x_0, x_2)$ and if the observation structure satisfies $X_0X_2 - X_1 - Y$, then any joint distribution $\widetilde{Q}(x_0, x_1, x_2) \triangleq \widetilde{Q}(x_0, x_2)\widehat{Q}(x_1)$ satisfying the information constraint

$$I_{\widetilde{Q}}(X_0; X_2) \leq I_{\widehat{Q}}(X_1; Y) \tag{62}$$

would be implementable; in particular, one may optimize $\widehat{Q}$ independently. In addition, if $w_1(x_1)$ is independent of x_1, the information constraint further simplifies as

$$I_{\widehat{Q}}(X_0;X_2) \leq \max_{\widehat{Q}} I_{\widehat{Q}}(X_1;Y), \tag{63}$$

and the implicit communication channel effectively becomes a dedicated channel.

We now characterize implementable distributions for the observation structure corresponding to case II in (5).

Theorem 14. *Consider the observation structure in case* II. *Let* $\overline{Q} \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ *be such that* $\sum_{(x_1,x_2)} \overline{Q}(x_0,x_1,x_2) = \rho_0(x_0)$ *for all* $x_0 \in \mathcal{X}_0$. *If the distribution* Q *defined as*

$$Q(x_0,x_1,x_2,y,u) = \mathrm{P}(u|x_0,x_1,x_2)\Gamma(y|x_0,x_1,x_2)\overline{Q}(x_0,x_1,x_2) \tag{64}$$

for $(x_0,x_1,x_2,y,u) \in \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y} \times \mathcal{U}$ *verifies the constraint*

$$I_Q(X_0;X_2) < I_Q(U;Y,X_2) - I_Q(U;X_0,X_2), \tag{65}$$

then $\overline{Q}$ *is implementable in the sense of Definition 2.*

Proof: Consider a joint distribution $Q_{X_0X_1X_2YU} \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y} \times \mathcal{U})$ that factorizes as in (64). As for the proof of Theorem 8, the idea is to construct the strategies $f : \underline{x}_0^{(b+1)}, \underline{x}_0^{(b)}, \underline{x}_2^{(b)} \mapsto \underline{x}_1^{(b)}$ and $g : \underline{y}^{(b)}, \underline{x}_2^{(b)} \mapsto \underline{x}_2^{(b+1)}$ from a block-Markov coding scheme that operates over B blocks. As illustrated in Fig. 3, Agent 1 will again communicate to Agent 2 during block b the actions that Agent 2 should play in block $b+1$; however, unlike the situation in Fig. 2, Agent 2 now only observes noisy versions of the actions of Agent 1. While coordination in case I relied on coding schemes similar to those for communication over state-dependent channel with state information available at the transmitter and receiver, coordination in case II will rely on Gel'fand-Pinsker schemes to deal with the absence of state information at the receiver.

We will again have to distinguish communication during block 1, which cannot be fully coordinated. We thus introduce another joint distribution $\widehat{Q}$ that factorizes as

$$\widehat{Q}(x_0,x_1,x_2,y,u) = \Gamma(y|x_0,x_1,x_2)P(u|x_0x_1x_2)\overline{Q}(x_1|x_0x_2)\widehat{Q}(x_0)\widehat{Q}(x_2), \tag{66}$$

which differs from $\overline{Q}$ by the independence of X_0 and X_2. Now, let $\epsilon > 0$. Let $R, R', \widehat{R}, \widehat{R}' > 0$, $\frac{\epsilon}{2} > \epsilon_3 > \epsilon_2 > \epsilon_1 > 0$ be real numbers and $n \in \mathbb{N}^*$ to be specified later. Define α as in (45) and B as in (46).

Source codebook generation for $b = 1$. Choose a sequence $\underline{x}_2^{(1)}(i_0) \in \mathcal{T}_{\epsilon_1}^n(Q_{X_2})$. This unique sequence is revealed to both agents.

Source codebooks generation for $b \in \{2,\dots,B+1\}$. Randomly and independently generate $\lceil 2^{nR} \rceil$ sequences according to $\Pi_{t=1}^n Q_{X_2}(x_{2,t})$. Label the sequences $\underline{x}_2^{(b)}(i)$ with $i \in \{1,\dots,2^{nR}\}$. These sequences constitute the source codebook $\mathcal{C}_s^{(b)}$, which is revealed to both agents.

Channel codebook generation for $b = 1$. Randomly and independently generate $2^{\alpha n(\widehat{R}'+\widehat{R})}$ sequences according to $\Pi_{t=1}^n \widehat{Q}_U(u_t)$. Label the sequences $\underline{u}^{(1)}(i,j)$ with $i \in \{1,\dots,2^{\alpha n\widehat{R}}\}$ and $j \in \{1,\dots,2^{\alpha n\widehat{R}'}\}$. These sequences constitute the channel codebook for block 1, which is revealed to both agents.

Channel codebook generation for $b \in \{2,\dots,B\}$. Randomly and independently generate $2^{n(R'+R)}$ sequences according to $\Pi_{t=1}^n Q_U(u_t)$. Label the sequences $\underline{u}^{(b)}(i,j)$ with $i \in \{1,\dots,2^{nR}\}$ and $j \in \{1,\dots,2^{nR'}\}$. These sequences constitute the channel codebook for block b, which is revealed to both agents.

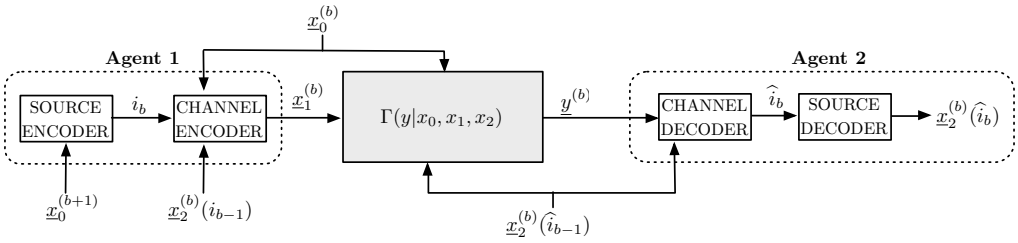

Fig. 3. Illustration of encoding and decoding operations in block b of the proposed block-coding scheme for case II.

Source encoding at Agent 1 in block $b \in \{1, \dots, B\}$. At the beginning of block b, Agent 1 uses its non-causal knowledge of the state $\underline{x}_0^{(b+1)}$ in the next block $b+1$ to look for an index i_b such that $(\underline{x}_0^{(b+1)}, \underline{x}_2^{(b+1)}(i_b)) \in \mathcal{T}_{\epsilon_1}^n(Q_{X_0X_2})$. If there is more than one such index, it chooses the smallest among them, otherwise it chooses $i_b = 1$.

Channel encoding at Agent in block $b = 1$. Agent 1 uses its knowledge of $(\underline{x}_0^{(1)}, \underline{x}_2^{(1)}(i_0))$ to look for an index j_1 such that

$$\left(\underline{u}^{(1)}(i_1, j_1), \underline{x}_0^{(1)}, \underline{x}_2^{(1)}(i_0)\right) \in \mathcal{T}_{\epsilon_2}^n(\widehat{Q}_{UX_0X_2}) \tag{67}$$

If there is more than one such index, it chooses the smallest among them, otherwise it chooses $j_1 = 1$. Finally, Agent 1 generates a sequence $\underline{x}_1^{(1)}$ by passing the sequences $\underline{u}^{(1)}(i_1, j_1)$, $\underline{x}_0^{(1)}$, and $\underline{x}_2^{(1)}(i_{b-1})$ through a DMC with transition probability $\widehat{Q}_{X_1|UX_0X_2}$; the sequence $\underline{x}_1^{(1)}$ is transmitted during block 1.

Channel encoding at Agent in block $b \in \{2, \dots, B\}$. Agent 1 uses its knowledge of $(\underline{x}_0^{(b)}, \underline{x}_2^{(b)}(i_{b-1}))$ to look for an index j_b such that

$$\left(\underline{u}^{(b)}(i_b, j_b), \underline{x}_0^{(b)}, \underline{x}_2^{(b)}(i_{b-1})\right) \in \mathcal{T}_{\epsilon_2}^n(Q_{UX_0X_2}) \tag{68}$$

If there is more than one such index, it chooses the smallest among them, otherwise it chooses $j_b = 1$. Finally, Agent 1 generates a sequence $\underline{x}_1^{(b)}$ by passing the sequences $\underline{u}^{(b)}(i_b, j_b)$, $\underline{x}_0^{(b)}$, and $\underline{x}_2^{(b)}(i_{b-1})$ through a DMC with transition probability $Q_{X_1|UX_0X_2}$; the sequence $\underline{x}_1^{(b)}$ is transmitted during block b.

Decoding at Agent 2 in block $b = 1$. At the end of block 1, Agent 2 observes the sequence of channel outputs $\underline{y}^{(1)}$ and knows its sequence of actions $(\underline{x}_2^{(1)}(\widehat{i}_{b-1}))$ in block 1. Agent 2 then looks for a pair of indices $(\widehat{i}_1, \widehat{j}_1)$ such that

$$\left(\underline{u}^{(b)}(\widehat{i}_1, \widehat{j}_1), \underline{y}^{(1)}, \underline{x}_2^{(1)}(i_0)\right) \in \mathcal{T}_{\epsilon_3}^n(\widehat{Q}_{UYX_2}). \tag{69}$$

If there is none or more than one such index, Agent 2 sets $\widehat{i}_1 = \widehat{j}_1 = 1$.

Channel decoding at Agent 2 in block $b \in \{2, \dots, B\}$. At the end of block b, Agent 2 observes the sequence of channel outputs $\underline{y}^{(b)}$ and knows its sequence of actions $(\underline{x}_2^{(b)}(\widehat{i}_{b-1}))$ in block b. Agent 2 then looks for a pair of indices $(\widehat{i}_b, \widehat{j}_b)$ such that

$$\left(\underline{u}^{(b)}(\widehat{i}_b, \widehat{j}_b), \underline{y}^{(b)}, \underline{x}_2^{(b)}(\widehat{i}_{b-1})\right) \in \mathcal{T}_{\epsilon_3}^n(Q_{UYX_2}). \tag{70}$$

If there is none or more than one such index, Agent 2 sets $\widehat{i}_b = \widehat{j}_b = 1$.

Source decoding at Agent 2 in block $b \in \{1, \dots, B\}$. Agent 2 transmits $\underline{x}_2^{(b)}(\widehat{i}_{b-1})$, where $\widehat{i}_{b-1}$ is its estimate of the message transmitted by Agent 1 in the previous block $b-1$, with the convention that $\widehat{i}_0 = i_0$.

Analysis. We follow the same approach as in the proof of Theorem 8. We define again the error event

$$E \triangleq \{(\underline{X}_0^T, \underline{X}_1^T, \underline{X}_2^T) \notin \mathcal{T}_\epsilon^n(\overline{Q})\} \tag{71}$$

and we proceed to show that $P(E)$ can be made arbitrarily small for n sufficiently large and a proper choice of the rates R, R', and $\widehat{R}$. We introduce the following events in each block $b \in \{1, \dots, B\}$.

$$\begin{aligned}
E_0 &\triangleq \{(I_1, J_1) \neq (\widehat{I}_1, \widehat{J}_1)\} \\
E_1^{(b)} &\triangleq \{(\underline{X}_0^{(b+1)}, \underline{x}_2^{(b)}(i)) \notin \mathcal{T}_{\epsilon_1}^n(Q_{X_0X_2}) \, \forall \, i \in \{1, \dots, 2^{nR}\}\} \\
E_2^{(b)} &\triangleq \{(\underline{u}^{(b)}(I_b, j), \underline{X}_0^{(b)}, \underline{x}_2^{(b)}(I_b)) \notin \mathcal{T}_{\epsilon_1}^n(Q_{UX_0X_2}) \, \forall \, j \, \{1, \dots, 2^{nR'}\}\} \\
E_3^{(b)} &\triangleq \left\{(\underline{u}^{(b)}(I_b, J_b), \underline{X}_0^{(b)}, X_1^{(b)}, \underline{x}_2^{(b)}(\widehat{I}_{b-1}), \underline{Y}^{(b)}) \notin \mathcal{T}_{\epsilon_3}^n(Q_{UX_0X_1X_2Y})\right\} \\
E_4^{(b)} &\triangleq \left\{((\underline{u}^{(b)}(i, j), \underline{x}_2^{(b)}(\widehat{I}_{b-1}), \underline{Y}^{(b)}) \in \mathcal{T}_{\epsilon_3}^n(Q) \text{ for some } (i, j) \neq (i_b, j_b)\right\}.
\end{aligned}$$

As in the proof of Theorem 8, $P(E)$ may be upper bounded as

$$\begin{aligned}
\mathrm{P}(E) \leq \mathrm{P}(E_0) + \sum_{b=1}^{B} \mathrm{P}(E_1^{(b)}) + \sum_{b=2}^{B} \mathrm{P}(E_2^{(b)} | E_1^{(b-1)c}) + \sum_{b=2}^{B} \mathrm{P}(E_4^{(b)} | E_1^{(b-1)c} \cap E_2^{(b-1)c} \cap E_3^{(b-1)c} \cap E_4^{(b-1)c} \cap E_0^c) \\
+ \sum_{b=2}^{B} \mathrm{P}(E_3^{(b)} | E_1^{(b-1)c} \cap E_2^{(b-1)c} \cap E_3^{(b-1)c} \cap E_4^{(b-1)c} \cap E_0^c)
\end{aligned} \tag{72}$$

As proved in Appendix D, the following lemmas show that all the averages over the random codebooks of the terms above vanish as $n \to \infty$.

Lemma 15. *If $\widehat{R} > I_{\widehat{Q}}(U; X_0X_2) + \delta(\epsilon_2)$ and $\widehat{R} + \widehat{R}' < I_{\widehat{Q}}(U; YX_2) - \delta(\epsilon_3)$, then*

$$\lim_{n\to\infty} E\left(\mathrm{P}(E_0)\right) = 0. \tag{73}$$

Lemma 16. *If $R > I_Q(X_0; X_2) + \delta(\epsilon_1)$, then for any $b \in \{1, \ldots, B\}$*

$$\lim_{n\to\infty} E\left(\mathrm{P}(E_1^{(b)})\right) = 0. \tag{74}$$

Lemma 17. *If $R' > I_Q(U; X_0, X_2) + \delta(\epsilon_2)$, then for any $b \in \{2, \ldots, B\}$*

$$\lim_{n\to\infty} E\left(\mathrm{P}(E_2^{(b)})|E_1^{(b-1)})\right) = 0. \tag{75}$$

Lemma 18. *For any $b \in \{2, \ldots, B\}$*

$$\lim_{n\to\infty} E\left(\mathrm{P}(E_3^{(b)}|E_1^{(b-1)c} \cap E_2^{(b-1)c} \cap E_3^{(b-1)c} \cap E_4^{(b-1)c} \cap E_0^{c})\right) = 0. \tag{76}$$

Lemma 19. *If $R + R' < I_Q(U; Y, X_2) - \delta(\epsilon_3)$, then for any $b \in \{2, \ldots, B\}$*

$$\lim_{n\to\infty} E\left(\mathrm{P}(E_4^{(b)}|E_1^{(b-1)c} \cap E_2^{(b-1)c} \cap E_3^{(b-1)c} \cap E_4^{(b-1)c} \cap E_0^{c})\right) = 0. \tag{77}$$

Therefore, if $I_{\widehat{Q}}(U; Y, X_2) - I_{\widehat{Q}}(U; X_0, X_2) > 0$ and $I_Q(X_0; X_2) < I_Q(U; Y, X_2) - I_Q(U; X_0, X_2)$, we can find ϵ_1, ϵ_2, and ϵ_3 small enough such that $\lim_{n\to\infty} E(\mathrm{P}(E)) = 0$. In particular, there must exists at least one sequence of codes such that $\lim_{n\to\infty} \mathrm{P}(E) = 0$. Since $\epsilon > 0$ can be chosen arbitrarily small, $\overline{Q}$ is an achievable empirical coordination. ∎

Although the proof of Theorem 14 relies on a separate source channel coding scheme, there is a key difference with respect to the usual Gel'fand-Pinsker coding scheme [22] and its extensions [35]. Indeed, while using the channel decoder's past outputs does not help in terms of improving the channel capacity, it is useful to achieve coordination. Specifically, a classical Gel'fand-Pinsker coding results would lead to an information constraint

$$I_Q(X_0; X_2) \le I_Q(U; Y) - I_Q(U; X_0, X_2) \tag{78}$$

which is more restrictive than (65).

By specializing to the case in which Agent 2 perfectly monitors Agent 1's actions, i.e., $Y = X_1$, equations (15), (43), (65), and (78) coincide with the information constraint $I_Q(X_0; X_2) \le H_Q(X_1|X_0, X_2)$. This shows that knowing X_0 strictly causally as in case I does not bring any performance improvement under perfect monitoring. Not that we have not proved whether the information constraint of Theorem 14 is a necessary condition for implementability in case II. One might be tempted to adopt a side information interpretation of the problem to derive the converse, since (65) resembles the situation of [39]; however, finding the appropriate auxiliary variables does not seem straightforward.

IV. Expected payoff optimization

In this section, we study the problem of determining the joint distribution(s) Q that leads to the maximal payoff in case I and case II. We establish two formulations of the problem: one that involves Q viewed as a function, and one that explicitly involves the vector of probability masses of Q. Although the latter form involves seemingly more complex notation, it is better suited to numerically determine the maximum expected payoff and turns out particularly useful in Section V. While we study the general optimization problem in Section IV-A, we focus on the case of perfect monitoring in Section IV-B, for which we are able to gain more insight into the structure of the optimal solution.

A. General optimization problem

From the results derived in Section III, the determination of the largest average payoff requires solving the following optimization problem, with $\ell \in \{\mathrm{I}, \mathrm{II}\}$:

$$\begin{array}{lll}
\text{minimize} & -\mathbb{E}_Q[w(X)] = -\displaystyle\sum_{x,y,u} Q(x,y,u)w(x) & \\
\text{s.t.} & -1 + \displaystyle\sum_{(x,y,u)} Q(x,y,u) & \overset{(c)}{=} 0 \\
\forall (x,y,u) \in \mathcal{X}\times\mathcal{Y}\times\mathcal{U}, & \dfrac{Q(x,y,u)}{\displaystyle\sum_{(y,u)} Q(x,y,u)} - \Gamma(y|x) & \overset{(d)}{=} 0 \\
\forall x_0 \in \mathcal{X}_0, & -\rho_0(x_0) + \displaystyle\sum_{x_1,x_2,y,u} Q(x,y,u) & \overset{(e)}{=} 0 \\
\forall (x,y,u) \in \mathcal{X}\times\mathcal{Y}\times\mathcal{U}, & -Q(x,y,u) & \overset{(f)}{\le} 0 \\
 & \Phi^{\ell}(Q) & \overset{(g)}{\le} 0
\end{array} \tag{79}$$

where $X \triangleq (X_0, X_1, X_2)$, $x \triangleq (x_0, x_1, x_2)$, and in case I, $\Phi^{\mathrm{I}}(\cdot)$ is defined in (34) while in case II

$$\Phi^{\mathrm{II}}(Q) \triangleq I_Q(X_0; X_2) - I_Q(U; Y, X_2) + I_Q(U; X_0, X_2). \tag{80}$$

A first aspect of the optimization problem to consider is its potential convexity [40]. Note that the objective function to be minimized is linear in Q. In addition, the constraints (c), (d), (e), and (f) restrict the domain to a convex subset of the unit simplex. Therefore, it suffices to show that the domain resulting from the additional constraint (g) is convex for the optimization problem to be convex. In case I, for which the set $\mathcal{U}$ reduces to a singleton, Lemma 7 already proves that Φ^{I} is a convex function of Q, which implies that the additional constraint (g) defines a convex domain. In case II, we have not proved that Φ^{II} is a convex function but, by using a time-sharing argument, it is always possible to make the domain convex. In the remaining of the paper, we assume this convexification is always performed, so that the optimization problem is again convex.

A second aspect to consider is whether Slater's condition holds, so that the Karush-Kuhn-Tucker (KKT) conditions become necessary conditions for optimality. Since the problem is convex, the KKT conditions would also be sufficient.

Proposition 20. *Slater's condition holds in cases I and II for irreducible channel transition probabilities i.e., such that* $\forall (x_0, x_1, x_2, y) \in \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}, \Gamma(y|x_0, x_1, x_2) > 0$.

Proof: We establish the existence of a strictly feasible point in case II, from which the existence for case I follows as a special case. Consider a distribution $Q(x_0, x_1, x_2, y, u)$ such that X_0, X_1, and X_2 are independent, and $U = X_1$. We can assume without loss of generality that the support of the marginals Q_{X_i}, $i \in \{0, 1, 2\}$ is full, i.e., $\forall x_i \in \mathcal{X}_i, Q_{X_i}(x_i) > 0$. If the channel transition probability is irreducible, note that $Q(x_0, x_1, x_2, y, u)$ is then strictly positive, making the constraint (f) inactive. As for inequality constraint (g), notice that

$$\begin{aligned} I_Q(X_0; X_2) - I_Q(U; Y, X_2) + I_Q(U; X_0, X_2) &= 0 - I_Q(X_1; Y) - I(X_1; X_2|Y) + I_Q(X_1; X_0, X_2) && (81)\\ &= -I_Q(X_1; Y) - I_Q(X_1; X_2|Y) && (82)\\ &= -H_Q(X_1) + H_Q(X_1|Y, X_2) && (83)\\ &< 0. && (84) \end{aligned}$$

Hence, the chosen distribution constitutes a strictly feasible point for the domain defined by constraints (c)-(g), and remains a strictly feasible point after convexification of the domain. ∎

Our objective is now to rewrite the above optimization problem more explicitly in terms of the vector of probability masses that describe Q. This is useful not only to exploit standard numerical solvers in Section V, but also to apply the KKT conditions in Section IV-B. We introduce the following notation. Without loss of generality, the finite sets $\mathcal{X}_k$ for $k \in \{0, 1, 2\}$ are written in the present section as set of indices $\mathcal{X}_k = \{1, \cdots, n_k\}$; similarly, we write $\mathcal{U} = \{1, \cdots, n_u\}$ and $\mathcal{Y} = \{1, \cdots, n_y\}$. With this convention, we define a bijective mapping $\psi^\ell : \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y} \times \mathcal{U} \to \{1, \cdots, n^\ell\}$ as

$$\psi^\ell(i', j', k', l', m') \triangleq m' + n_u(l' - 1) + n_u n_y(k' - 1) + n_u n_y n_2(j' - 1) + n_u n_y n_2 n_1(i' - 1), \tag{85}$$

which maps a realization $(i', j', k', l', m') \in \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y} \times \mathcal{U}$ to a unique index $\psi^\ell(i', j', k', l', m') \in \{1, \cdots, n^\ell\}$. We also set $n^{\mathrm{I}} \triangleq n_0 n_1 n_2 n_y$ and $n^{\mathrm{II}} \triangleq n_0 n_1 n_2 n_y n_u$. This allows us to introduce the vector of probability masses $q^{n^\ell} = (q_1, q_2, \cdots, q_{n^\ell})$ for $\ell \in \{\mathrm{I}, \mathrm{II}\}$, in which each component q_i, $i \in \{1, \cdots, n^\ell\}$, is equal to $Q((\psi^\ell)^{-1}(i))$, and the vector of payoff values $w^{n^\ell} = (w_1, w_2, \cdots, w_{n^\ell}) \in \mathbb{R}^{n^\ell}$, in which each component w_i is the payoff of $(\psi^\ell)^{-1}(i)$. The relation between the mapping Q (resp. w) and the vector q^{n^ℓ} (resp. w^{n^ℓ}) is summarized in Table II.

Using the proposed indexing scheme, the optimization problem is written in standard form as follows.

$$\begin{array}{lll} \text{minimize} & -\mathbb{E}_Q[w(X)] = -\displaystyle\sum_{i=1}^{n^\ell} q_i w_i & \\ \text{s.t.} & -1 + \displaystyle\sum_{i=1}^{n^\ell} q_i & \overset{(h)}{=} 0 \\ \forall i \in \{1, \ldots, n^\ell\}, & \dfrac{q_i}{\Theta_i} - \Gamma_i & \overset{(i)}{=} 0 \\ \forall i \in \{1, \ldots, n_0\}, & -\rho_0(i) + \displaystyle\sum_{j=1+(i-1)n_1 n_2 n_y n_u}^{i n_1 n_2 n_y n_u} q_j & \overset{(j)}{=} 0 \\ \forall i \in \{1, \cdots, n^\ell\}, & -q_i & \overset{(k)}{\leq} 0 \\ & \phi^\ell(q^{n^\ell}) & \overset{(\ell)}{\leq} 0 \end{array} \tag{86}$$

TABLE II
CHOSEN INDEXATION FOR THE PAYOFF VECTOR w AND PROBABILITY DISTRIBUTION VECTOR q. BOLD LINES DELINEATE BLOCKS OF SIZE $n_1n_2n_yn_u$ AND EACH BLOCK CORRESPONDS TO A GIVEN VALUE OF THE RANDOM STATE X_0. THE $5-$UPLETS ARE SORTED ACCORDING TO A LEXICOGRAPHIC ORDER.

Index of q_i	X_0	X_1	X_2	Y	U
1	1	1	1	1	1
2	1	1	1	1	2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
n_u	1	1	1	1	n_u
n_u+1	1	1	1	2	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$2n_u$	1	1	1	2	n_u
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n_1n_2n_yn_u-n_u+1$	1	n_1	n_2	n_y	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n_1n_2n_yn_u$	1	n_1	n_2	n_y	n_u
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(n_0-1)n_1n_2n_yn_u+1$	n_0	1	1	1	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n_0n_1n_2n_yn_u$	n_0	n_1	n_2	n_y	n_u

where

$$\Theta_i \stackrel{\Delta}{=} \sum_{\substack{j\in\{1,\dots,n_yn_u\}\\ k\in\{1,\dots,n_0n_1n_2\}}} q_{(k-1)n_yn_u+j}\cdot\mathbb{1}_{\{(k-1)n_yn_u\le i\le kn_yn_u-1\}} \tag{87}$$

and $\forall i\in\{1,\cdots,n_0\}$, ρ_{0i} corresponds to $\mathrm{P}(X_0=i)=\rho_0(i)$ and $\forall i\in\{1,\cdots,n^\ell\}$, Γ_i corresponds to the value of $\Gamma(y|x_0,x_1,x_2)$, according to Table II. As for the function associated with inequality constraint (ℓ), it writes in case II (case I follows by specialization with $|\mathcal{U}|=1$) as follows:

$$\begin{aligned}\phi^{\mathrm{II}}(q^{n^{\mathrm{II}}}) &= I_{q^{n^{\mathrm{II}}}}(X_0;X_2)-I_{q^{n^{\mathrm{II}}}}(U;Y,X_2)+I_{q^{n^{\mathrm{II}}}}(U;X_0,X_2)\\ &= H_{q^{n^{\mathrm{II}}}}(X_0)-H_{q^{n^{\mathrm{II}}}}(U,X_0|X_2)+H_{q^{n^{\mathrm{II}}}}(U|Y,X_2)\\ &= H_{q^{n^{\mathrm{II}}}}(X_0)+H_{q^{n^{\mathrm{II}}}}(X_2)-H_{q^{n^{\mathrm{II}}}}(X_0,X_2,U)+H_{q^{n^{\mathrm{II}}}}(X_2,Y,U)-H_{q^{n^{\mathrm{II}}}}(X_2,Y)\end{aligned} \tag{88}$$

with

$$H_{q^{n^{\mathrm{II}}}}(X_0)=-\sum_{i=1}^{n_0}\left[\Big(\sum_{j=1+(i-1)n_1n_2n_yn_u}^{in_1n_2n_yn_u} q_j\Big)\log\Big(\sum_{j=1+(i-1)n_1n_2n_yn_u}^{in_1n_2n_yn_u} q_j\Big)\right], \tag{89}$$

$$H_{q^{n^{\mathrm{II}}}}(X_2)=-\sum_{i=1}^{n_2}\left[\Big(\sum_{j=1}^{n_0n_1}\sum_{k=1}^{n_yn_u} q_{(i-1)n_un_y+(j-1)n_2n_yn_u+k}\Big)\log\Big(\sum_{j=1}^{n_0n_1}\sum_{k=1}^{n_yn_u} q_{(i-1)n_un_y+(j-1)n_2n_yn_u+k}\Big)\right], \tag{90}$$

$$H_{q^{n^{\mathrm{II}}}}(X_2,Y,U)=-\sum_{i=1}^{n_2n_yn_u}\left[\Big(\sum_{j=1}^{n_0n_1} q_{(j-1)n_2n_yn_u+i}\Big)\log\Big(\sum_{j=1}^{n_0n_1} q_{(j-1)n_2n_yn_u+i}\Big)\right], \tag{91}$$

$$\begin{aligned}H_{q^{n^{\mathrm{II}}}}(X_0,X_2,U)=-\sum_{i=1}^{n_0}\sum_{j=1}^{n_2}\sum_{k=1}^{n_u}\Bigg[&\Big(\sum_{l=1}^{n_0n_1}\sum_{m=1}^{n_y} q_{(i-1)n_1n_2n_yn_u+(j-1)n_yn_u+k+(l-1)n_2n_yn_u+(m-1)n_u}\Big)\\ &\log\Big(\sum_{l=1}^{n_0n_1}\sum_{m=1}^{n_y} q_{(i-1)n_1n_2n_yn_u+(j-1)n_yn_u+k+(l-1)n_2n_yn_u+(m-1)n_u}\Big)\Bigg],\end{aligned} \tag{92}$$

and

$$H_{q^{n^{\mathrm{II}}}}(X_2,Y)=-\sum_{i=1}^{n_2n_y}\left[(\sum_{j=1}^{n_0n_1}\sum_{k=1}^{n_u}q_{(j-1)n_2n_yn_u+(i-1)n_u+k})\log(\sum_{j=1}^{n_0n_1}\sum_{k=1}^{n_u}q_{(j-1)n_2n_yn_u+(i-1)n_u+k})\right]. \tag{93}$$

This formulation is directly exploited in Section IV-B and in Section V.

B. Optimization problem for perfect monitoring

In the case of perfect monitoring, for which agent 2 perfectly monitors Agent 1's actions and $Y=X_1$, the information constraints (43) and (65) coincide and

$$\phi(q^n)\triangleq\phi^{\mathrm{I}}(q^{n^{\mathrm{I}}})=\phi^{\mathrm{II}}(q^{n^{\mathrm{II}}})=H_{q^n}(X_2)-H_{q^n}(X_2|X_0)-H_{q^n}(X_1|X_0,X_2) \tag{94}$$

with $q^n=(q_1,\cdots,q_n)$, $n=n_0n_1n_2$. To further analyze the relationship between the vector of payoff values w^n and an optimal joint distribution q^n, we explicitly express the KKT conditions. The Lagrangian is

$$\mathcal{L}(q^n,\lambda^n,\mu_0,\mu^{n_0},\lambda_{\mathrm{IC}})=-\sum_{i=1}^{n}w_iq_i+\lambda_iq_i+\mu_0\left[-1+\sum_{i=1}^{n}q_i\right]+\sum_{j=1}^{n_0}\mu_j\left[-\rho_{0i}+\sum_{i=1+(j-1)n_1n_2}^{jn_1n_2}q_i\right]+\lambda_{\mathrm{IC}}\phi(q^n) \tag{95}$$

where $\lambda^n=(\lambda_1,\cdots,\lambda_n)$, $\mu^{n_0}=(\mu_1,\cdots,\mu_{n_0})$, and the subscript IC stands for information constraint.

A necessary and sufficient condition for a distribution q^n to be an optimum point is that it is a solution of the following system:

$$\forall\, i\in\{1,\ldots,n\},\ \frac{\partial\mathcal{L}}{\partial q_i}=-w_i-\lambda_i+\mu_0+\sum_{j=1}^{n_0}\mu_j\mathbb{1}_{\{1+n_1n_2(j-1)\le i\le jn_1n_2\}}+\lambda_{\mathrm{IC}}\frac{\partial\phi}{\partial q_i}(q^n)=0 \tag{96}$$

$$(h),(i),(j) \tag{97}$$

$$\forall\, i\in\{1,\ldots,n\},\ \lambda_i\ge 0 \tag{98}$$

$$\lambda_{\mathrm{IC}}\ge 0 \tag{99}$$

$$\forall\, i\in\{1,\ldots,n\},\ \lambda_iq_i=0 \tag{100}$$

$$\lambda_{\mathrm{IC}}\phi(q^n)=0 \tag{101}$$

where

$$\forall i\in\{1,\ldots,n\},\ \frac{\partial\phi}{\partial q_i}(q^n)=\left[-\sum_{k=1}^{n_0}\left(\mathbb{1}_{\{1+(k-1)n_1n_2\le i\le kn_1n_2\}}\log\sum_{j=1+(k-1)n_1n_2}^{kn_1n_2}q_j\right)\right.$$
$$\left.-\sum_{k=1}^{n_2}\mathbb{1}_{\{i\in\{k,k+n_2,\ldots,k+(n_0n_1-1)n_2\}\}}\log\sum_{j=0}^{n_0n_1-1}q_{k+jn_2}-1+\log q_i\right]. \tag{102}$$

In the following, we assume that there exists a permutation of $\{1,\cdots,n\}$ such that the vector of payoff values w^n after permutation of the components is strictly ordered. A couple of observations can then be made by inspecting the KKT conditions above. First, if the expected payoff were only maximized under the constraints (h) and (k), the best joint distribution would be to only assign probability to the greatest element of the vector w^n; in other words the best q^n would correspond to a vertex of the unit simplex $\Delta(\mathcal{X}_0\times\mathcal{X}_1\times\mathcal{X}_2)$. However, as the distribution of the random state fixed by constraint (j), at least n_0 components of q^n should be positive. It is readily verified that under constraints $(h),(j)$, and (k), the optimal solution is that for each x_0 the optimal pair (x_1,x_2) is chosen; therefore, q^n possesses exactly n_0 positive components. This corresponds to the costless communication scenario. Now, in the presence of the additional information constraint (ℓ), the optimal solutions contain in general more than n_0 positive components. This is due to the fact that optimal communication schemes between the two agents requires several symbols of $\mathcal{X}_1$ to be associated with a given realization of the state. In fact, as shown in the following proposition, there is a unique optimal solution under the assumptions made.

Proposition 21. *If there exists a permutation such that the payoff vector w^n is strictly ordered, then the optimization problem (86) has a unique solution.*

Proof: Assume $\lambda_{\mathrm{IC}}=0$ in the Lagrangian, i.e., $\phi(q^n)<0$ and the information constraint is active for optimal solutions. Further assume that a candidate solution of the optimization problem q^n has two or more positive components in a block of size n_1n_2 associated with a given realization x_0 (see Table II). Then, there exist two indices (i_1,i_2) such that $\lambda_{i_1}=0$ and $\lambda_{i_2}=0$. Consequently, the conditions on the gradient $\frac{\partial\mathcal{L}}{\partial q_i}=0$ for $i\in\{i_1,i_2\}$ imply that $w_{i_1}=w_{i_2}$, which contradicts the assumption of w^n being strictly ordered under permutation. Therefore, a candidate solution only possesses a single positive

component per block associated with a given realization x_0, which means that X_1 and X_2 are deterministic functions of X_0. Hence, $H_{q^n}(X_2|X_0) = H_{q^n}(X_1|X_0X_2) = 0$ and the information constraint reads $H_{q^n}(X_2) < 0$, which is impossible. Hence, $\lambda_{\rm IC} > 0$.

From Lemma 7, we know that $\phi(q^n)$ is strictly convex. Since $\lambda_{\rm IC} > 0$ the Lagrangian is then the sum of linear functions and a strictly convex function. Since it is also continuous and optimized over a compact and convex set, there exists a maximum point and it is unique. ■

Apart from assuming that w^n can be strictly ordered, Proposition 21 does not assume anything on the values of the components of w^n. In practice, for a specific problem it will be relevant to exploit the special features of the problem of interest to better characterize the relationship between the payoff function (which is represented by w^n) and the optimal joint probability distributions (which are represented by the vector q^n). This is one of the purposes of the next section.

V. Coded power control

We now exploit the framework developed in the previous sections to study the problem of power control in interference networks. In this context, the agents are the transmitters and the random state corresponds to the global channel state, i.e., all the channel gains associated with the different links between transmitters and receivers. *Coded power control (CPC)* consists in embedding information about the global channel state into transmit power levels themselves rather than using a dedicated signaling channel. Provided that the power levels of a given transmitter can be observed by the other transmitters, the sequence of power levels can be used to coordinate with the other transmitters. Typical mechanisms through which agents may observe power levels include sensing, as in cognitive radio settings, or feedback, as often assumed in interference networks. One of the salient features of coded power control is that interference is directly managed in the radio-frequency domain and does not require baseband detection or decoding, which is very useful in systems such as heterogeneous networks. The main goal of this section is to assess the limiting performance of coded power control and its potential performance gains over other approaches, such as the Nash equilibrium power control policies of a given non-cooperative game. The corresponding comparison is especially relevant since conventional distributed power control algorithms such as the iterative water-filling algorithm do not exploit the opportunity to exchange information through power levels or vectors to implement a better solution (which e.g., Pareto-dominates the Nash equilibrium power control policies).

A. Coded power control over interference channels

We consider an interference channel with two transmitters and two receivers. We specialize the model to the multiple-access channel in Section V-E, for which we develop and analyze an explicit non-trivial power control code.

By denoting g_{ij} the channel gain between Transmitter i and Receiver j, each realization of the global channel state is given by

$$x_0 = (g_{11}, g_{12}, g_{21}, g_{22}), \tag{103}$$

where $g_{ij} \in \mathcal{G}$, $|\mathcal{G}| < \infty$; it is further assumed that the channel gains g_{ij} are independent. Each alphabet $\mathcal{X}_i, |\mathcal{X}_i| < \infty$, $i \in \{1,2\}$, represents the set of possible power levels for Transmitter i. Assuming that the sets are discrete is of practical interest, as there exist wireless communication standards in which the power can only be decreased or increased by step and in which quantized channel state information is used. In addition, the use of discrete power levels may not induce any loss of optimality [41] w.r.t. the continuous case, and we elaborate more on this point in Section V-B.

We consider three stage payoff functions $w^{\rm rate}$, $w^{\rm SINR}$, and $w^{\rm energy}$, which respectively represent the sum-rate, the sum-SINR, and the sum-energy efficiency. Specifically,

$$\begin{aligned} w^{\rm rate}: \quad \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 &\to \mathbb{R}_+ \\ (x_0, x_1, x_2) &\mapsto \sum_{i=1}^{2} \log_2\left(1 + \underbrace{\frac{g_{ii}x_i}{\sigma^2 + g_{-ii}x_{-i}}}_{\mathrm{SINR}_i}\right), \end{aligned} \tag{104}$$

$$\begin{aligned} w^{\rm SINR}: \quad \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 &\to \mathbb{R}_+ \\ (x_0, x_1, x_2) &\mapsto \sum_{i=1}^{2} \frac{g_{ii}x_i}{\sigma^2 + g_{-ii}x_{-i}}, \end{aligned} \tag{105}$$

$$\begin{aligned} w^{\rm energy}: \quad \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 &\to \mathbb{R}_+ \\ (x_0, x_1, x_2) &\mapsto \sum_{i=1}^{2} \frac{F\left(1 + \frac{g_{ii}x_i}{\sigma^2 + g_{-ii}x_{-i}}\right)}{x_i} \end{aligned} \tag{106}$$

where $\mathcal{X}_0 = \mathcal{G}^4$; the notation $-i$ stands for the transmitter other than i; σ^2 corresponds to the reception noise level; $F : \mathbb{R}_+ \to [0,1]$ is a sigmoidal and increasing function which typically represents the block success rate, see e.g., [42], [43]. The function F is chosen so that $w^{\rm energy}$ is continuous and has a limit when $x_i \to 0$.

Finally, we consider the following three possible observation structures.

- *Perfect monitoring*, in which Agent 2 directly observes the actions of Agent 1, i.e., $Y = X_1$;
- *BSC monitoring*, in which Agent 2 observes the actions of Agent 1 through a binary symmetric channel (BSC), i.e., $Y = X_1 \oplus Z_1$ and X_1, Y, and Z_1 are Bernoulli random variables and $\oplus$ is the $2-$modulo addition;
- *Noisy SINR feedback monitoring*, in which Agent 2 observes a noisy version of the SINR of Agent 1 as illustrated in Fig. 4. This corresponds to a scenario in which a feedback channel exists between Receiver 2 and Transmitter 2.

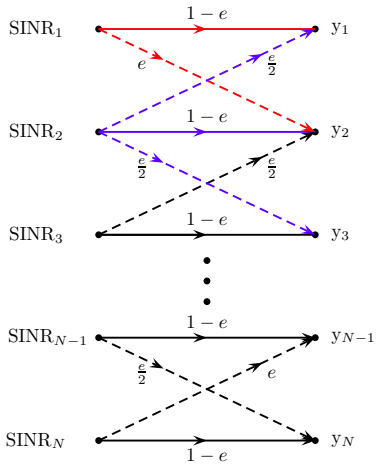

Fig. 4. The signal observed by Transmitter 2 lies in an $N-$symbol alphabet i.e., $|\mathcal{Y}| = N$. The symbols correspond to possible values for the SINR at the receiver which is associated with Transmitter 2. Transmitter 2 observes the actual value of its SINR with probability $1 - e$ while there is probability $e \geq 0$ that a neighboring SINR is observed. In the simulations, $e = 0.1$.

The performance of coded power control will be assessed against that of the following three benchmark power control policies.

- *Nash-equilibrium power control (NPC) policy*. In such a policy, each transmitter aims at maximizing an individual stage payoff function $u_i(x_0, x_1, x_2)$. In the sum-rate, sum-SINR, and sum-energy efficiency cases, these individual stage payoff-functions are respectively given by $u_i(x_0, x_1, x_2) = \log_2(1 + \mathrm{SINR}_i)$, $u_i(x_0, x_1, x_2) = \mathrm{SINR}_i$, and $u_i(x_0, x_1, x_2) = \frac{F(\mathrm{SINR}_i)}{x_i}$. In the sum-rate and sum-SINR cases, the unique Nash equilibrium is $(x_1^{\mathrm{NE}}, x_2^{\mathrm{NE}}) = (P_{\max}, P_{\max})$, irrespectively of the value of x_0, $P_{\max}$ being the maximal power level for the transmitters. In the sum-energy efficiency case, the unique non-trivial Nash equilibrium may be determined numerically and generally requires some knowledge of x_0, depending on how it is implemented.
- *Semi-coordinated power control (SPC) policy*. This policy corresponds to a basic coordination scheme in which Transmitter 1 optimizes its power knowing that Transmitter 2 transmits at full power; SPC requires the knowledge of the current state realization at Transmitter 1. Specifically, $x_2 = P_{\max}$, $x_1^\dagger \in \arg\max_{x_1} w^{\mathrm{r}}(x_0, x_1, P_{\max})$, $\mathrm{r} \in \{\mathrm{rate}, \mathrm{SINR}, \mathrm{energy}\}$.
- *Costless-communication power control (CCPC) policy.* This policy corresponds to the situation in which transmitters may communicate at not cost, so that they may jointly optimize their powers to achieve the maximum of the payoff function w^{r} at every stage. In such a case there is no information constraint, and the performance of CCPC provides an upper bound for the performance of all other policies.

The communication signal-to-noise ratio (SNR) is defined as

$$\mathrm{SNR(dB)} \triangleq 10 \log_{10} \frac{P_{\max}}{\sigma^2}. \tag{107}$$

B. Influence of the payoff function

The objective of this subsection is to numerically assess the relative performance gain of CPC over SPC in the case of perfect monitoring. We assume that the channel gains $g_{ij} \in \{g_{\min}, g_{\max}\}$ are Bernoulli distributed $g_{ij} \sim \mathcal{B}(p_{ij})$ with $p_{ij} \triangleq \mathrm{P}(g_{ij} = g_{\min})$; with our definition of X_0 in (103), this implies that $|\mathcal{X}_0| = 16$. All numerical results in this subsection

are obtained for $g_{\min} = 0.1$, $g_{\max} = 2$ and $(p_{11}, p_{12}, p_{21}, p_{22}) = (0.5, 0.1, 0.1, 0.5)$. The sets of transmit powers $\mathcal{X}_1$, $\mathcal{X}_2$ are both assumed to be the same alphabet of size four $\{P_1, P_2, P_3, P_4\}$, with $P_1 = 0, P_2 = \frac{P_{\max}}{3}, P_3 = \frac{2P_{\max}}{3}, P_4 = P_{\max}$. The quantity $P_{\max}$ is given by the operating SNR and $\sigma^2 = 1$. The function F is chosen as a typical instance of the efficiency function used in [43], i.e.,

$$F(x) = \exp\left(-\frac{2^{0.9} - 1}{x}\right). \tag{108}$$

For all $\mathrm{r} \in \{\mathrm{rate}, \mathrm{SINR}, \mathrm{energy}\}$, the relative performance gain with respect to the SPC policy is

$$\left(\frac{\mathbb{E}_{\overline{Q}^\star}(w^{\mathrm{r}})}{\mathbb{E}_{\rho_0}(\max\limits_{x_1} w(x_0, x_1, P_{\max}))} - 1\right) \times 100 \tag{109}$$

where $\overline{Q}^\star$ is obtained by solving optimization problem (86) under perfect monitoring. This optimization is numerically performed using the Matlab function fmincon. Fig. 5 illustrates the relative performance gain in % w.r.t. the SPC policy for the sum-energy efficiency, while Fig. 6 illustrates it for the sum-SINR and sum-rate.

As shown in Fig. 5, our simulation results suggest that CPC provides significant performance gains for the sum-energy efficiency. This may not be surprising, as the payoff function (106) is particularly sensitive to the lack of coordination; in fact, as the transmit power becomes high, $\frac{F(\mathrm{SINR}_i)}{x_i} \to \frac{1}{x_i}$, which means that energy efficiency decreases rapidly. As shown in Fig. 6, the performance gains of CPC for the sum-SINR and the sum-rate are more moderate, with gains as high as 43% for the sum-SINR and 25% for the sum-rate; nevertheless, such gains are still significant, and would be larger if we used NPC instead of SPC as the reference case, as often done in the literature of distributed power control.

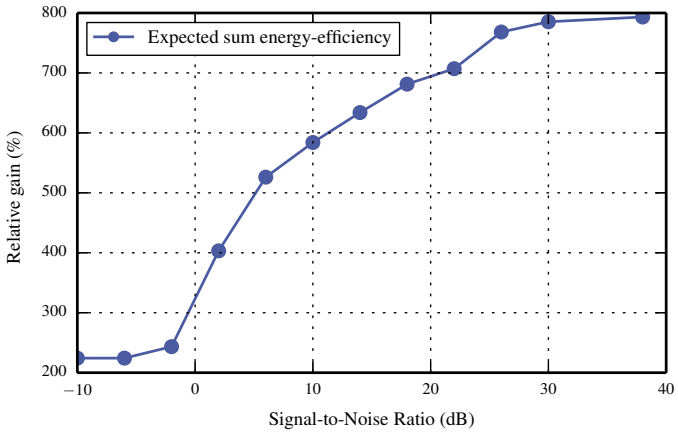

Fig. 5. Relative sum-energy gain of coded power control (CPC) with perfect monitoring over semi-coordinated power control (SPC).

We conclude this subsection by providing the marginals $\overline{Q}^\star_{X_1}(x_1) = \sum\limits_{x_0, x_2} \overline{Q}^\star(x_0, x_1, x_2)$, $\overline{Q}^\star_{X_2}(x_2) = \sum\limits_{x_0, x_1} \overline{Q}^\star(x_0, x_1, x_2)$, and joint distribution $\overline{Q}^\star_{X_1 X_2}(x_1, x_2) = \sum\limits_{x_0} \overline{Q}^\star(x_0, x_1, x_2)$ of the optimal joint distribution for CPC and CCPC in Table III and Table IV, respectively. In both cases, the results correspond to the maximization of the sum-rate payoff function w^{rate} and $\mathrm{SNR} = 10$ dB. Table III shows that, without information constraint, the sum-rate is maximized when the transmitters correlate their power levels so that only three pairs of transmit power levels are used out of 16. This result is consistent with [41], which proves that, for interference channels with 2 transmitter-receiver pairs, there is no loss of optimality in terms of w^{rate} by operating over a binary set $\{0, P_{\max}\}$ instead of a continuous interval $[0, P_{\max}]$. Interestingly, as seen in Table IV, the three best configurations of the CCPC policy are exploited $44.3 + 42.9 + 2.1 = 89.3\%$ of the time in the CPC policy, despite the presence of communication constraints between the two transmitters.

C. Influence of the observation structure

In this subsection, we focus on the observation structure defined by case I in (4) and we restrict our attention to the sum-rate payoff function w^{rate}. The set of powers is restricted to a binary set $\mathcal{X}_1 = \mathcal{X}_2 = \{0, P_{\max}\}$, but unlike the study in Section

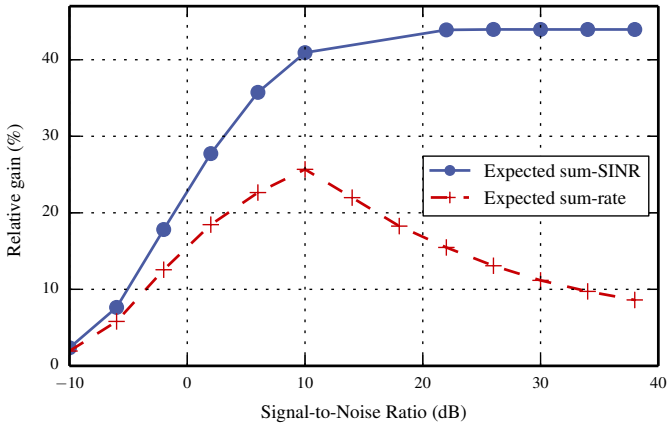

Fig. 6. Relative sum-SINR gain and sum-rate gain of coded power control (CPC) with perfect monitoring over semi-coordinated power control (SPC).

TABLE III
OPTIMAL MARGINAL AND JOINT DISTRIBUTIONS (EXPRESSED IN %) FOR THE SUM-RATE PAYOFF FUNCTION OF THE CCPC POLICY, WITH SNR $= 10$ DB AND WITH FOUR POSSIBLE TRANSMIT POWER LEVELS $\{0, \frac{10}{3}, \frac{20}{3}, 10\}$.

$(\overline{Q}^{\star}_{X_1}(x_1), \overline{Q}^{\star}_{X_2}(x_2), \overline{\mathbf{Q}}^{\star}_{\mathbf{X_1X_2}}(\mathbf{x_1}, \mathbf{x_2}))$ in %	$x_1 = 0$	$x_1 = \frac{10}{3}$	$x_1 = \frac{20}{3}$	$x_1 = 10$
$x_2 = 0$	(47.5,47.5,**00.0**)	(00.0,47.5,**00.0**)	(00.0,47.5,**00.0**)	(52.5,47.5,**47.5**)
$x_2 = \frac{10}{3}$	(47.5,00.0,**00.0**)	(00.0,00.0,**00.0**)	(00.0,00.0,**00.0**)	(52.5,00.0,**00.0**)
$x_2 = \frac{20}{3}$	(47.5,00.0,**00.0**)	(00.0,00.0,**00.0**)	(00.0,00.0,**00.0**)	(52.5,00.0,**00.0**)
$x_2 = 10$	(47.5,52.5,**47.5**)	(00.0,52.5,**00.0**)	(00.0,52.5,**00.0**)	(52.5,52.5,**05.5**)

TABLE IV
OPTIMAL MARGINAL AND JOINT DISTRIBUTIONS (EXPRESSED IN %) FOR THE SUM-RATE PAYOFF FUNCTION OF THE CPC POLICY, WITH SNR $= 10$ DB AND WITH FOUR POSSIBLE TRANSMIT POWER LEVELS $\{0, \frac{10}{3}, \frac{20}{3}, 10\}$.

$(\overline{Q}^{\star}_{X_1}(x_1), \overline{Q}^{\star}_{X_2}(x_2), \overline{\mathbf{Q}}^{\star}_{\mathbf{X_1X_2}}(\mathbf{x_1}, \mathbf{x_2}))$ in %	$x_1 = 0$	$x_1 = \frac{10}{3}$	$x_1 = \frac{20}{3}$	$x_1 = 10$
$x_2 = 0$	(44.4,50.4,**00.1**)	(02.6,50.4,**00.9**)	(08.0,50.4,**06.5**)	(45.0,50.4,**42.9**)
$x_2 = \frac{10}{3}$	(44.4,00.0,**00.0**)	(02.6,00.0,**00.0**)	(08.0,00.0,**00.0**)	(45.0,00.0,**00.0**)
$x_2 = \frac{20}{3}$	(44.4,00.0,**00.0**)	(02.6,00.0,**00.0**)	(08.0,00.0,**00.0**)	(45.0,00.0,**00.0**)
$x_2 = 10$	(44.4,49.6,**44.3**)	(02.6,49.6,**01.7**)	(08.0,49.6,**01.5**)	(45.0,49.6,**02.1**)

V-B, we do not limit ourselves to perfect monitoring. Fig. 7 shows the relative performance gain w.r.t. the SPC policy as a function of SNR for three different observation structures. The performance of CPC for BSC monitoring is obtained assuming a probability of error of 5%, i.e., $Z_1 \sim \mathcal{B}(0.05)$, $\mathrm{P}(Z_1 = 1) = 0.05$. The performance of CPC for noisy SINR feedback monitoring is obtained assuming $e = 0.1$; in this case, it can be checked that the SINR can take one of $N = 7$ distinct values.

Fig. 7 suggests that CPC provides a significant performance gain over SPC over a wide range of operating SNRs irrespective of the observation structure. Interestingly, for SNR $= 10$ dB, the relative gain of CPC only drops from 22% with perfect monitoring to 18% with BSC monitoring, which suggest that for observation structures with typical noise levels the benefits of CPC are somewhat robust to observation noise. Similar observations can be made for SINR feedback monitoring. Note that one would obtain higher performance gains by considering NPC as the reference policy or by considering scenarios with stronger interference.

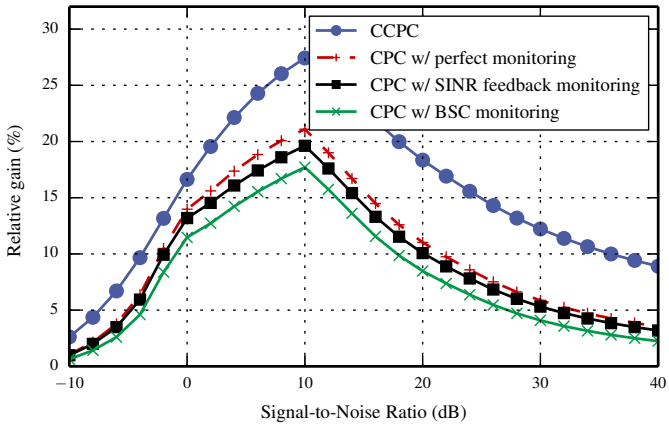

Fig. 7. Relative sum-rate gain of costless communication power control (CCPC) and coded power control (CPC) over semi-coordinated power control (SPC) under various monitoring assumptions in the observation structure of Case I.

D. Influence of the state knowledge

In this subsection, we restrict our attention to CPC with BSC monitoring with the same parameters as in Section V-C, but we consider both Case I and Case II defined in (4) and (5), respectively. The results for Case II are obtained assuming that $|\mathcal{U}| = 10$. While we already know that the performance of CPC is the same in Case I and Case II with perfect monitoring, the results in Fig. 8 suggest that, for typical values of the observation noise, not knowing the past realizations of the global channel state at Transmitter 2 only induces a small performance loss.

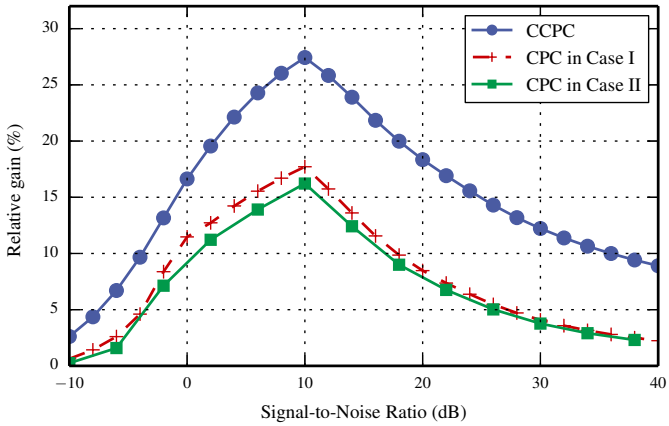

Fig. 8. Relative sum-rate gain of coded power control (CPC) and costless communication power control (CCPC) over semi-coordinated power control (SPC) for the binary symmetric channel (BSC) monitoring in the observation structure of Case I and Case II.

E. Influence of the coordination scheme

In this last subsection, we assess the benefits of CPC for an explicit code that operates over blocks of length $n = 3$. To simplify the analysis and clarify the interpretation, several assumptions are made. First, we consider a multiple-access channel,

which is a special case of the interference channel studied earlier with two transmitters and a single receiver, so that the global channel state comprises only two components (g_1, g_2). Second, we assume that the global channel state X_0 takes values in the binary alphabet $\mathcal{X}_0 \in \{(g_{\min}, g_{\max}), (g_{\max}, g_{\min})\}$, and is distributed according to Bernoulli random variable $\mathcal{B}(p)$ with $p = \mathrm{P}\big(X_0 = (g_{\min}, g_{\max})\big)$. In the remaining of this subsection, we identify the realization $(g_{\min}, g_{\max})$ with "0" and $(g_{\max}, g_{\min})$ wth "1", so that we may write $\mathcal{X}_0 = \{0, 1\}$. Third, we assume that the transmitters may only choose power values in $\{P_{\min}, P_{\max}\}$, and we identify power $P_{\min}$ with "0" and power $P_{\max}$ with "1", so that we may also write $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$. Finally, we consider the case of perfect monitoring and we restrict our attention to the sum-SINR payoff function w^{SINR}.

The values of the payoff function used in numerical simulations are provided in Fig. 9 as the entries in a matrix. Each matrix corresponds to a different choice of the state x_0; in each matrix, the choice of the row corresponds to the action x_1 of Transmitter 1, which the choice of the column corresponds to the action x_2 of Transmitter 2.

	$P_{\min}$	$P_{\max}$
$P_{\min}$	0	20
$P_{\max}$	1	$\simeq 10$

$x_0 = (g_{\min}, g_{\max})$

	$P_{\min}$	$P_{\max}$
$P_{\min}$	0	1
$P_{\max}$	20	$\simeq 10$

$x_0 = (g_{\max}, g_{\min})$

Fig. 9. Payoff matrix of w^{SINR} for power control over multiple access channel. Numerical values of the payoff correpond to $g_{\min} = 0.1$, $g_{\max} = 2$, $\sigma^2 = 1$, $P_{\min} = 0$, $P_{\max} = 10$.

The coordination code of length 3 that we develop next can be seen as a separate source channel code, which consists of a source code with distortion and a channel code with side information. The source encoder and decoder are defined by the mappings

$$f_{\mathrm{S}} : \begin{array}{ccc} \mathcal{X}_0^3 & \to & \{m_0, m_1\} \\ \underline{x}_0 & \mapsto & i \end{array}, \tag{110}$$

$$g_{\mathrm{S}} : \begin{array}{ccc} \{m_0, m_1\} & \to \mathcal{X}_2^3 & \\ i & \mapsto & \underline{x}_2 \end{array}. \tag{111}$$

Note that the chosen source code only uses 2 messages $\{m_0, m_1\}$ to represent the 8 possible sequences $\underline{x}_0$. The exact choice of f_{S} and g_{S} is provided after we describe the channel code.

In each block b, Transmitter 1's channel encoder implements the mapping

$$\underline{x}_1^{(b)} = f_{\mathrm{C}}(\underline{x}_0^{(b)}, \underline{x}_2^{(b)}, i_{b+1}) \tag{112}$$

where $i_{b+1} = f_{\mathrm{S}}(\underline{x}_0^{(b+1)})$ is the index associated with the sequence $\underline{x}_2^{(b+1)}$. The idea behind the design of the channel encoder f_{C} is the following. If Transmitter 1 did not have to transmit the index i_{b+1}, its optimal encoding would be to exploit its knowledge of $(x_0^3(b), x_2^3(b))$ to choose the sequence $x_1^3(b)$ resulting in the highest average payoff in block b. However, to communicate the index i_{b+1}, Transmitter 1 will instead choose to transmit the sequence $x_1^3(b)$ with the highest average payoff in block b if $i_{b+1} = m_0$, or the sequence $\underline{x}_1^{(b)}$ with the second highest average payoff in block b if $i_{b+1} = m_1$. Note that Transmitter 2 is able to perfectly decode this encoding given its knowledge of $\underline{x}_0^{(b)}$, $\underline{x}_2^{(b)}$, and $\underline{x}_1^{(b)}$ at the end of block b. Formally, f_{C} is defined as follows. The sequence $\underline{x}_1$ is chosen as

$$\underline{x}_1 = \underline{x}_1' \oplus \underline{d} \tag{113}$$

where the modulo-two addition is performed component-wise,

$$\underline{x}_1' = \arg\max_{\underline{x}_1 \in \mathcal{X}_1^3} \sum_{t=1}^{3} w^{\mathrm{SINR}}(x_{0,t}, x_{1,t}, x_{2,t}), \tag{114}$$

$$\underline{d} = (0, 0, 0) \text{ if } i_{b+1} = m_0, \tag{115}$$

$$d^3 = \arg\max_{\underline{d} \text{ s.t. } \Omega(\underline{d})=1} \sum_{t=1}^{3} w^{\mathrm{SINR}}(x_{0,t}, x_{1,t}' \oplus d_t, x_{2,t}) \text{ if } i_{b+1} = m_1 \tag{116}$$

where Ω is the Hamming weight function that is, the number of ones in the sequence $\underline{d} \in \{0, 1\}^3$. If the argmax set is not a singleton set, we choose the sequence with the smallest Hamming weight.

To complete the construction, we must specify how the source code is designed. Here, we choose the mappings f_{S} and g_{S} that maximize the expected payoff $\mathbb{E}(w^{\mathrm{SINR}})$ knowing the operation of the channel code. The source code resulting from an

TABLE V
PROPOSED SOURCE CODING AND DECODING FOR $p = \frac{1}{2}$.

x_0^3	Index $i = f_{\mathrm{S}}(x_0^3)$	$g_{\mathrm{S}}(i)$
(0,0,0)	m_0	(1,1,1)
(0,0,1)	m_0	(1,1,1)
(0,1,0)	m_0	(1,1,1)
(0,1,1)	m_0	(1,1,1)
(1,0,0)	m_0	(1,1,1)
(1,0,1)	m_0	(1,1,1)
(1,1,0)	m_1	(0,0,1)
(1,1,1)	m_1	(0,0,1)

TABLE VI
PROPOSED CHANNEL CODING FOR $p = \frac{1}{2}$.

$x_0^3(b)$	$x_2^3(b)$	i_{b+1}	$x_1^3(b)$
(0,0,0)	(1,1,1)	m_0	(0,0,0)
		m_1	(0,0,1)
(0,0,1)	(1,1,1)	m_0	(0,0,1)
		m_1	(0,0,0)
(0,1,0)	(1,1,1)	m_0	(0,1,0)
		m_1	(0,0,0)
(0,1,1)	(1,1,1)	m_0	(0,1,1)
		m_1	(0,0,1)
(1,0,0)	(1,1,1)	m_0	(1,0,0)
		m_1	(0,0,0)
(1,0,1)	(1,1,1)	m_0	(1,0,1)
		m_1	(0,0,1)
(1,1,0)	(0,0,1)	m_0	(1,1,0)
		m_1	(1,1,1)
(1,1,1)	(0,0,1)	m_0	(1,1,1)
		m_1	(1,1,0)

exhaustive search is given in Table V, and the corresponding channel code is given in Table VI. The detailed expression of the expected payoff required for the search is provided in Appendix E.

The proposed codes admit to an intuitive interpretation. For instance, the first line of Table VI indicates that if the channel is bad for Transmitter 1 for the three stages of block b, then Transmitter 1 remains silent over the three stages of the block while Transmitter 2 transmits at all three stages. In contrast, the last line of Table VI shows that if the channel is good for Transmitter 1 for the three stages of block b, then Transmitter 1 transmit at all stages while Transmitter 2 remains silent two thirds of the time. While this is suboptimal for this specific global state realization, this is required to allow coordination and average optimality of the code.

To conclude this section, we compare the performance of this short code with the best possible performance that would be obtained with infinitely long codes. As illustrated in Fig. 10, while the performance of the short code suffers from a small penalty compared to that of ideal codes with infinite block length, it still offers a significant gain w.r.t. the SPC policy and it outperforms the NPC policy.

VI. CONCLUSION

In this paper, we adopted the view that distributed control policies or resource allocation policies in a network are joint source-channel codes. Essentially, an agent of a distributed network may convey its knowledge of the network state by *encoding* it into a sequence of actions, which can then be *decoded* by the agents observing that sequence. As explicitly shown in Section V-E, the purpose of such "coordination codes" is neither to convey information reliably nor to meet a requirement in terms of maximal distortion level, but to provide a high expected payoff. Consequently, coordination codes must implement a trade-off between sending information about the future realizations of the network state, which plays the role of an information source and is required to coordinate future actions, and achieving an acceptable payoff for the current state of the network. Considering

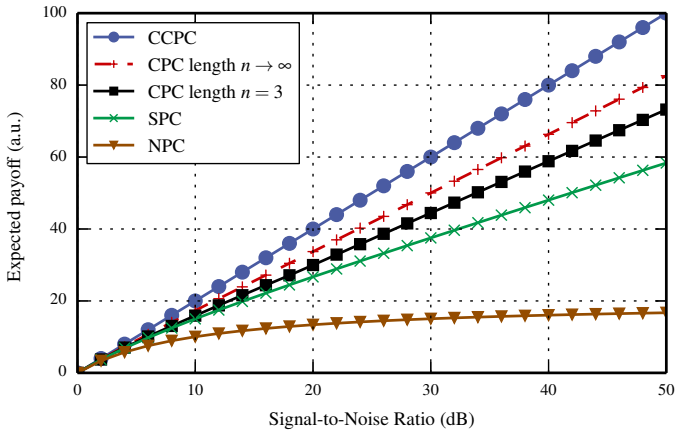

Fig. 10. Expected payoff versus SNR for different power control policies.

the large variety of payoff functions in control and resource allocation problems, an interesting issue is whether universal codes performing well within classes of payoff functions can be designed.

Remarkably, since a distributed control policy or resource allocation policy is interpreted as a code, Shannon theory naturally appears to measure the efficiency of such policies. While the focus of this paper was limited to a small network of two agents, the proposed methodology to derive the best coordination performance in a distributed network is much more general. The assumptions made in this paper are likely to be unsuited to some application scenarios, but provide encouraging preliminary results to further research in this direction. For example, as mentioned in Section I, a detailed comparison between coded power control and iterative water-filling like algorithms would lead to consider a symmetric observation structure while only an asymmetric structure is studied in this paper. The methodology to assess the performance of good coded policies consists in deriving the right information constraint(s) by building the proof on Shannon theory for the problem of multi-source coding with distortion over multi-user channels wide side information and then to use this constraint to find an information-constrained maximum of the payoff (common payoff case) or the set of Nash equilibrium points which are compatible with the constraint (non-cooperative game case). As a key observation of this paper, the observation structure of a multi-person decision-making problem corresponds in fact to a multiuser channel. Therefore, multi-terminal Shannon theory is not only relevant for pure communication problems but also for any multi-person decision-making problem. The above observation also opens new challenges for Shannon-theorists since decision-making problems define new communication scenarios.

APPENDIX A
PROOF OF PROPOSITION 5

Assume $\overline{Q}$ is an achievable empirical coordination. Then, for any $\epsilon > 0$,

$$\begin{aligned} ||\mathbb{E}(\mathrm{T}_{X^T}) - \overline{Q}||_1 &\leq \mathbb{E}(||\mathrm{T}_{X^T} - \overline{Q}||_1) && (117)\\ &= \mathbb{E}(||\mathrm{T}_{X^T} - \overline{Q}||_1 \mid ||\mathrm{T}_{X^T} - \overline{Q}||_1 \geq \epsilon)\mathrm{P}(||\mathrm{T}_{X^T} - \overline{Q}||_1 \geq \epsilon) \\ &\quad + \mathbb{E}(||\mathrm{T}_{X^T} - \overline{Q}||_1 \mid ||\mathrm{T}_{X^T} - \overline{Q}||_1 < \epsilon)\mathrm{P}(||\mathrm{T}_{X^T} - \overline{Q}||_1 < \epsilon) && (118)\\ &\leq 2\mathrm{P}(||\mathrm{T}_{X^T} - \overline{Q}||_1 \geq \epsilon) + \epsilon. && (119) \end{aligned}$$

Hence, $\forall \epsilon > 0 \quad \lim_{T\to\infty} ||\mathbb{E}(\mathrm{T}_{X^T}) - \overline{Q}||_1 \leq \epsilon$, which means that $\overline{Q}$ is implementable.

APPENDIX B
PROOF OF LEMMA 7

The function $\Phi = \Phi^{\mathrm{I}}$ can be rewritten as $\Phi(Q) = H_Q(X_0) - H_Q(Y, X_0|X_2) + H_Q(Y|X_0, X_2, X_1)$. The first term $H_Q(X_0) = -\sum_{x_0} \rho_0(x_0) \log \rho_0(x_0)$ is a constant w.r.t. Q. The third term is linear w.r.t. Q since, with Γ fixed,

$$H_Q(Y|X_0, X_2, X_1) = -\sum_{x_0,x_1,x_2,y} Q(x_0, x_1, x_2, y) \log \mathrm{P}(y|x_0, x_1, x_2) \tag{120}$$

$$= -\sum_{x_0,x_1,x_2,y} Q(x_0, x_1, x_2, y) \log \Gamma(y|x_0, x_1, x_2). \tag{121}$$

It is therefore sufficient to prove that $H_Q(Y, X_0|X_2)$ is concave. Let $\lambda_1 \in [0, 1]$, $\lambda_2 = 1-\lambda_1$, $(Q_1, Q_2) \in \Delta^2(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})$ and $Q = \lambda_1 Q_1 + \lambda_2 Q_2$. By using the standard notation $Z^0 = \emptyset$, $Z^n = (Z_1, \cdots, Z_n)$, Z being a generic sequence of random variables, we have that:

$$H_Q(Y, X_0|X_2) = -\sum_{x_0,x_2,y} \left(\sum_{x_1,i} \lambda_i Q_i(x_0, x_1, x_2, y) \right) \log \left[\frac{\sum_{x_1,i} \lambda_i Q_i(x_0, x_1, x_2, y)}{\sum_i \lambda_i P_{X_2}^{Q_i}(x_2)} \right] \tag{122}$$

$$= -\sum_{x_0,x_2,y} \left(\sum_i \lambda_i \sum_{x_1} Q_i(x_0, x_1, x_2, y) \right) \log \left[\frac{\sum_i \lambda_i \sum_{x_1} Q_i(x_0, x_1, x_2, y)}{\sum_i \lambda_i P_{X_2}^{Q_i}(x_2)} \right] \tag{123}$$

$$> -\sum_i \lambda_i \sum_{x_0,x_2,y} \left(\sum_{x_1} Q_i(x_0, x_1, x_2, y) \right) \log \left[\frac{\lambda_i \sum_{x_1} Q_i(x_0, x_1, x_2, y)}{\lambda_i P_{X_2}^{Q_i}(x_2)} \right] \tag{124}$$

$$= -\sum_i \lambda_i \sum_{x_0,x_2,y} \left(\sum_{x_1} Q_i(x_0, x_1, x_2, y) \right) \log \left[\frac{\sum_{x_1} Q_i(x_0, x_1, x_2, y)}{P_{X_2}^{Q_i}(x_2)} \right] \tag{125}$$

$$= \lambda_1 H_{Q_1}(Y, X_0|X_2) + \lambda_2 H_{Q_2}(Y, X_0|X_2) \tag{126}$$

where the strict inequality comes from the log sum inequality [28], with:

$$a_i = \lambda_i Q^i(x_0, x_1, x_2) \tag{127}$$

and

$$b_i = \lambda_i Q_{X_2}^i(x_2) \tag{128}$$

for $i \in \{1, 2\}$ and for all (x_0, x_1, x_2) such that $Q_{X_2}^i(x_2) > 0$.

APPENDIX C
LEMMAS USED IN PROOF OF THEOREM 8

A. Proof of Lemma 9

Recall that $T = \alpha n + (B-1)n$ with our coding scheme. Note that

$$\|T_{\underline{x}_0^T \underline{x}_1^T \underline{x}_2^T} - \overline{Q}\|_1 = \sum_{x_0,x_1,x_2} \left| \sum_{t=1}^{T} \frac{1}{T} \mathbb{1}_{\{(x_{0,t}, x_{1,t}, x_{2,t}) = (x_0, x_1, x_2)\}} - \overline{Q}(x_0, x_1, x_2) \right| \tag{129}$$

$$= \sum_{x_0,x_1,x_2} \left| \sum_{t=1}^{\alpha n} \frac{1}{T} \mathbb{1}_{\left\{(x_{0,t}^{(1)}, x_{1,t}^{(1)}, x_{2,t}^{(1)}) = (x_0, x_1, x_2)\right\}} + \sum_{b=2}^{B} \sum_{t=1}^{n} \frac{1}{T} \mathbb{1}_{\left\{(x_{0,t}^{(b)}, x_{1,t}^{(b)}, x_{2,t}^{(b)}) = (x_0, x_1, x_2)\right\}} - \overline{Q}(x_0, x_1, x_2) \right| \tag{130}$$

$$\leq \frac{\alpha n}{T} \sum_{x_0,x_1,x_2} \left| \sum_{t=1}^{\alpha n} \frac{1}{\alpha n} \mathbb{1}_{\left\{(x_{0,t}^{(1)}, x_{1,t}^{(1)}, x_{2,t}^{(1)}) = (x_0, x_1, x_2)\right\}} - \overline{Q}(x_0, x_1, x_2) \right|$$
$$+ \frac{n(B-1)}{T} \frac{1}{B-1} \sum_{b=2}^{B} \sum_{x_0,x_1,x_2} \left| \sum_{t=1}^{n} \frac{1}{n} \mathbb{1}_{\left\{(x_{0,t}^{(b)}, x_{1,t}^{(b)}, x_{2,t}^{(b)}) = (x_0, x_1, x_2)\right\}} - \overline{Q}(x_0, x_1, x_2) \right| \tag{131}$$

$$\leq \frac{2\alpha}{B-1+\alpha} + \frac{1}{B-1} \sum_{b=2}^{B} \|T_{\underline{x}_0^{(b)} \underline{x}_1^{(b)} \underline{x}_2^{(b)}} - \overline{Q}\|_1, \tag{132}$$

where (131) follows from the sub-additivity of $\|\cdot\|_1$ and (132) follows from $\frac{n(B-1)}{T} \leq 1$ and $\|P - Q\|_1 \leq 2$ for $P, Q \in \Delta(\mathcal{X})$.

B. Proof of Lemma 10

Since $\underline{x}_2^{(1)}(i_0)$ and $\underline{x}_0^{(1)}$ are known to the transmitter and the receiver, the channel can multiplexed and demultiplexed based on the realizations $x_{2,t}^{(1)}$ and $x_{0,t}^{(1)}$ for $t \in \{1, \cdots, \alpha n\}$. Since $\underline{x}_0^{(1)}$ is drawn according to the i.i.d. distribution Q_{X_0} and $\underline{x}_2^{(1)}(i_0) \in \mathcal{T}_{\epsilon_1}^n(Q_{X_2})$, then by the conditional typicality lemma $(\underline{x}_0^{(1)}, \underline{x}_2^{(1)}(i_0)) \in \mathcal{T}_{\epsilon_2}^n(\widehat{Q}_{X_0X_2})$ with overwhleming probability as $n \to \infty$. Following [31, Section 7.4], the probability of decoding error goes to zero as $n \to \infty$ if $\hat{R} < (1-\epsilon_2) I_{\widehat{Q}}(X_1; Y|X_0X_2) - \delta(\epsilon_2)$.

C. Proof of Lemma 11

Note that

$$\mathbb{E}\left(\mathrm{P}(E_1^{(b)})\right) = \mathrm{P}\left((\underline{X}_0^{(b+1)}, \underline{X}_2^{(b)}(i)) \notin \mathcal{T}_{\epsilon_1}^n(Q_{X_0X_2}) \text{ for all } i \in \{1, \ldots, 2^{nR}\}\right) \tag{133}$$

with $\underline{X}_0^{(b)}$ distributed according to $\prod_{t=1}^n Q_{X_0}$, and $\underline{X}_2^{(b)}(i)$ independent of each other distributed according to $\prod_{t=1}^n Q_{X_2}$. Hence, the result directly follows from the covering lemma [31, Lemma 3.3], with the following choice of parameters.

$$U \leftarrow \emptyset \qquad X^n \leftarrow \underline{X}_0^{(b+1)} \qquad \hat{X}^n(m) \text{ with } m \in \mathcal{A} \leftarrow \underline{X}_2^{(b)}(i) \text{ with } i \in \{1, \ldots, 2^{nR}\}.$$

D. Proof of Lemma 12

The result follows from a careful application of the conditional typicality lemma. Note that conditioning on $E_1^{(b-1)c}$ ensures that $(\underline{X}_0^{(b)}, \underline{x}_2^{(b)}(I_{b-1})) \in \mathcal{T}_{\epsilon_1}^n(Q_{X_0X_2})$, while conditioning on $E_2^{(b-1)c} \cap E_3^{(b-1)c} \cap E_0^c$ guarantees that $\widehat{I}_{b-1} = I_{b-1}$. Consequently,

$$\begin{aligned}
&\mathrm{P}\left(E_2^{(b)} | E_1^{(b-1)c} \cap E_2^{(b-1)c} \cap E_3^{(b-1)c} \cap E_0^c\right) \\
&\quad = \mathrm{P}\Big((\underline{X}_0^{(b)}, \underline{x}_1^{(b)}(I_b, \underline{X}_0^{(b)}, \widehat{I}_{b-1}), \underline{x}_2^{(b)}(\widehat{I}_{b-1}), \underline{Y}^{(b)}) \notin \mathcal{T}_{\epsilon_2}^n(Q_{X_0X_2X_1Y}) \\
&\qquad\qquad |(\underline{X}_0^{(b)}, \underline{x}_2^{(b)}(I_{b-1})) \in \mathcal{T}_{\epsilon_1}^n(Q_{X_0X_2}) \cap \widehat{I}_{b-1} = I_{b-1}\Big) \qquad (134) \\
&\quad = \sum_{i,j,\underline{x}_0} p_{\widehat{I}_{b-1}, I_b, \underline{X}_0^{(b)}}(i, j, \underline{x}_0) \sum_{\underline{y}} \Gamma(\underline{y}|\underline{x}_0, \underline{x}_1^{(b)}(j, \underline{x}_0, i), \underline{x}_2^{(b)}(i)) \mathbb{1}_{\left\{(\underline{x}_0^{(b)}, \underline{x}_1^{(b)}(j, \underline{x}_0, i), \underline{x}_2^{(b)}(i), \underline{y}) \notin \mathcal{T}_{\epsilon_1}^n(Q_{X_0X_1X_2Y})\right\}}, \qquad (135)
\end{aligned}$$

where $p_{\widehat{I}_{b-1}, I_b, \underline{X}_0^{(b)}}$ denotes the joint distribution of $\widehat{I}_{b-1}, I_b, \underline{X}_0^{(b)}$ given $\underline{X}_0^{(b)}, \underline{X}_2(I_{b-1})) \in \mathcal{T}_{\epsilon_1}^n(Q_{X_0X_2})$ and $\widehat{I}_{b-1} = I_{b-1}$. Upon taking the average over the random codebooks, we obtain

$$\begin{aligned}
&\mathbb{E}\left(\mathrm{P}\left(E_2^{(b)} | E_1^{(b-1)c} \cap E_2^{(b-1)c} \cap E_3^{(b-1)c} \cap E_0^c\right)\right) \\
&\qquad = \sum_{i,j,\underline{x}_0} \mathbb{E}\left(p_{\widehat{I}_{b-1}, I_b, \underline{X}_0^{(b)}}(i, j, \underline{x}_0) \mathbb{E}\left(\sum_{\underline{y}} \Gamma(\underline{y}|\underline{x}_0, \underline{X}_1^{(b)}(j, \underline{x}_0, i), \underline{X}_2^{(b)}(i)) \right.\right. \\
&\qquad\qquad \left.\left. \mathbb{1}_{\left\{(\underline{x}_0^{(b)}, \underline{X}_1^{(b)}(j, \underline{x}_0, i), \underline{X}_2^{(b)}(i), \underline{y}) \notin \mathcal{T}_{\epsilon_1}^n(Q_{X_0X_1X_2Y})\right\}} | \underline{X}_2^{(b)}(i)) \right)\right). \qquad (136)
\end{aligned}$$

The inner expectation is therefore

$$\mathrm{P}\left((\underline{x}_0, \underline{X}_1, \underline{x}_2, \underline{Y}) \notin \mathcal{T}_{\epsilon_1}^n(Q_{X_0X_1X_2Y})\right), \tag{137}$$

where $(\underline{X}_1, \underline{Y})$ is distributed according to $\prod_{t=1}^n Q_{X_1|X_0X_2}(x_{1,t}|x_{2,t}x_{0,t}) \Gamma(y_t|x_{0,t}x_{1,t}x_{2,t})$, and $(\underline{x}_0, \underline{x}_2) \in \mathcal{T}_{\epsilon_1}^n(Q_{X_0X_2})$. The conditional typicality lemma [31, p. 27] guarantees that (137) vanishes as $n \to \infty$.

E. Proof of Lemma 13

The result is a consequence of the packing lemma. Note that

$$\begin{aligned}
&\mathrm{P}\left(E_3^{(b)} | E_1^{(b-1)c} \cap E_2^{(b-1)c} \cap E_3^{(b-1)c} \cap E_0^c\right) = \mathrm{P}\Big((\underline{X}_0^{(b)}, \underline{x}_1^{(b)}(i, \underline{X}_0^{(b)}, \widehat{I}_{b-1}), \underline{x}_2^{(b)}(\widehat{I}_{b-1}), \underline{Y}^{(b)}) \in \mathcal{T}_{\epsilon_2}^n(Q) \text{ for some } i \neq I_b \\
&\qquad\qquad |(\underline{X}_0^{(b)}, \underline{x}_2^{(b)}(I_{b-1})) \in \mathcal{T}_{\epsilon_1}^n(Q_{X_0X_2}) \cap \widehat{I}_{b-1} = I_{b-1}\Big) \qquad (138) \\
&= \sum_{k,j,\underline{x}_0} p_{\widehat{I}_{b-1}, I_b, \underline{X}_0^{(b)}}(k, j, \underline{x}_0) \sum_{\underline{y}} \Gamma(\underline{y}|\underline{x}_0, \underline{x}_1^{(b)}(j, \underline{x}_0, k), \underline{x}_2^{(b)}(k)) \mathbb{1}_{\left\{(\underline{x}_0^{(b)}, \underline{x}_1^{(b)}(i, \underline{x}_0, k), \underline{x}_2^{(b)}(k), \underline{y}) \in \mathcal{T}_{\epsilon_1}^n(Q_{X_0X_1X_2Y}) \text{ for some } i \neq j\right\}}, \qquad (139)
\end{aligned}$$

Hence,

$$\mathbb{E}\left(\mathrm{P}\left(E_3^{(b)}|E_1^{(b-1)c}\cap E_2^{(b-1)c}\cap E_3^{(b-1)c}\cap E_0^c\right)\right) \tag{140}$$

$$= \mathbb{E}\left(\sum_{k,j,\underline{x}_0} p_{\widehat{I}_{b-1},I_b,\underline{X}_0^{(b)}}(k,j,\underline{x}_0)\mathbb{E}\left(\sum_{\underline{y}}\Gamma(\underline{y}|\underline{x}_0,\underline{X}_1^{(b)}(j,\underline{x}_0,k),\underline{X}_2^{(b)}(k))\right.\right.$$

$$\left.\left.\mathbb{1}_{\left\{(\underline{x}_0^{(b)},\underline{X}_1^{(b)}(i,\underline{x}_0,k),\underline{X}_2^{(b)}(k),\underline{y})\in\mathcal{T}_{\epsilon_1}^n(Q_{X_0X_1X_2Y}) \text{ for some } i\neq j\right\}}\right)|\underline{X}_2^{(b)}(k)\right) \tag{141}$$

$$= \sum_j p_{I_b}(j)\mathrm{P}\left((\underline{X}_0,\underline{X}_1^{(b)}(i),\underline{X}_2^{(b)},\underline{Y})\in\mathcal{T}_{\epsilon_1}^n(Q_{X_0X_1X_2Y}) \text{ for some } i\neq I_b|I_b=j\right). \tag{142}$$

Note that $|\{i : i \neq i_b\}| = 2^{nR} - 1$, and the random sequences $\underline{X}_1^{(b)}(i)$ are generated according to $\prod_{i=1}^n Q_{X_1|X_0X_2}$ and are conditionally independent of $\underline{Y}$ given $(\underline{X}_0, \underline{X}_2)$. Hence, from [31, Lemma 3.1], we know that if $R < I_Q(X_1;Y|X_0X_2) - \delta(\epsilon_2)$, then $\mathbb{E}(P(E_3))$ vanishes as $n \to \infty$.

Appendix D
Lemmas used in proof of Theorem 14

A. Proof of Lemma 15

The proof of this result is similar to that of the following Lemmas, using the "uncoordinated" distribution $\widehat{Q}$ instead of Q. For brevity, we omit the proof.

B. Proof of Lemma 16

The result follows from the covering lemma as in the proof of Lemma 11 given in Appendix C.

C. Proof of Lemma 17

Note that

$$\mathbb{E}\left(\mathrm{P}(E_2^{(b)}|E_1^{(b-1)}) = \mathrm{P}(\underline{U}^{(b)}(I_b,j),\underline{X}_0^{(b)},\underline{X}_2^{(b)}(I_{b-1}))\notin\mathcal{T}_{\epsilon_2}^n(Q_{UX_0X_2})\right), \tag{143}$$

where $(\underline{X}_0^{(b)}, \underline{X}_2^{(b)}(I_{b-1})) \in \mathcal{T}_{\epsilon_1}^n(Q_{X_0X_2})$, and $\underline{U}^{(b)}(I_b, j)$ are generated independently of each other according to $\prod_{t=1}^n Q_U$. Hence, the result follows directly from the covering lemma [31, Lemma 3.3] with the following choice of parameters.

$$U \leftarrow \emptyset \qquad X^n \leftarrow \underline{X}_0^{(b)}, \underline{X}_2^{(b)}(I_{b-1}) \qquad \hat{X}^n(m) \text{ with } m \in \mathcal{A} \leftarrow \underline{U}^{(b)}(I_b, j) \text{ with } j \in \{1, \ldots, 2^{nR'}\}.$$

D. Proof of Lemma 18

As in the proof of Lemma 12, the result follows from a careful application of the conditional typicality lemma. We omit the details, which are essentially the same as those in Appendix C-D, with the necessary modifications to introduce the codewords $\underline{u}^{(b)}$.

E. Proof of Lemma 19

The results follows from an analysis essentially identical to that of Appendix C-E with $\underline{U}$ in place of $\underline{X}_1$.

Appendix E
Expression of the expected payoff W_3 which allows the best mappings f_S and g_S to be selected

We introduce the composite mapping $\chi_S = g_S \circ f_S$. For the channel code defined in Section V-E, the expected payoff only depends on the mappings f_S and χ_S, and we denote it by $W_3(f_S, \chi_S)$. The following notation is used below: $\chi_S(x_0^3) = (\chi_1(x_0^3), \chi_2(x_0^3), \chi_3(x_0^3))$ to stand for the three components of χ_S.

It can be checked that

$$W_3(f_S,\chi_S) = \sum_{(i,j,k)\in\{0,1\}^3} W_{ijk}(f_S,\chi_S) \tag{144}$$

where:

$$W_{000}(f_{\mathrm{S}},\chi_{\mathrm{S}}) = \underbrace{p^3}_{\mathrm{P}[x_0^3[b]=(0,0,0)]}\Bigg[\underbrace{P_0(f_{\mathrm{S}})}_{\mathrm{P}[i_{b+1}=m_0]} \times \underbrace{\Bigg(\max_{x_1^3\in\{0,1\}^3}\{\sum_{t=1}^{3} w(0,x_{1,t},\chi_t((0,0,0)))\}\Bigg)}_{\text{Best payoff}} + (1-P_0(f_{\mathrm{S}})) \times \underbrace{\Bigg(\max_{x_1^3\in\{0,1\}^3\setminus\mathcal{X}_1^{000}}\{\sum_{t=1}^{3} w(0,x_{1,t},\chi_t((0,0,0)))\}\Bigg)}_{\text{Second best payoff}}\Bigg], \tag{145}$$

$$\mathcal{X}_1^{000} = \arg\max_{x_1^3\in\{0,1\}^3}\{\sum_{t=1}^{3} w(0,x_{1,t},\chi_t((0,0,0)))\}, \tag{146}$$

$$W_{001}(f_{\mathrm{S}},\chi_{\mathrm{S}}) = p^2(1-p)\Bigg[P_0(f_{\mathrm{S}}) \times \Bigg(\max_{x_1^3\in\{0,1\}^3}\Big\{w(0,x_{1,1},\chi_1((0,0,1))) + w(0,x_{1,2},\chi_2((0,0,1))) + w(1,x_{1,3},\chi_3((0,0,1)))\Big\}\Bigg) + (1-P_0(f_{\mathrm{S}})) \times \Big(\max_{x_1^3\in\{0,1\}^3\setminus\mathcal{X}_1^{001}}\Big\{w(0,x_{1,1},\chi_1((0,0,1))) + w(0,x_{1,2},\chi_2((0,0,1))) + w(1,x_{1,3},\chi_3((0,0,1)))\Big\}\Big)\Bigg], \tag{147}$$

$$\mathcal{X}_1^{001} = \arg\max_{x_1^3\in\{0,1\}^3}\{w(0,x_{1,1},\chi_1((0,0,1))) + w(0,x_{1,2},\chi_2((0,0,1))) + w(1,x_{1,3},\chi_3((0,0,1)))\}, \tag{148}$$

$$W_{010}(f_{\mathrm{S}},\chi_{\mathrm{S}}) = p^2(1-p)\Bigg[P_0(f_{\mathrm{S}}) \times \Bigg(\max_{x_1^3\in\{0,1\}^3}\Big\{w(0,x_{1,1},\chi_1(0,1,0)) + w(1,x_{1,2},\chi_2(0,1,0)) + w(0,x_{1,3},\chi_3(0,1,0))\Big\}\Bigg) + (1-P_0(f_{\mathrm{S}})) \times \Big(\max_{x_1^3\in\{0,1\}^3\setminus\mathcal{X}_1^{010}}\Big\{w(0,x_{1,1},\chi_1(0,1,0)) + w(1,x_{1,2},\chi_2(0,1,0)) + w(0,x_{1,3},\chi_3(0,1,0))\Big\}\Big)\Bigg], \tag{149}$$

$$\mathcal{X}_1^{010} = \arg\max_{x_1^3\in\{0,1\}^3}\{w(0,x_{1,1},\chi_1(0,1,0)) + w(1,x_{1,2},\chi_2(0,1,0)) + w(0,x_{1,3},\chi_3(0,1,0))\}, \tag{150}$$

$$W_{100}(f_{\mathrm{S}},\chi_{\mathrm{S}}) = p^2(1-p)\Bigg[P_0(f_{\mathrm{S}}) \times \Bigg(\max_{x_1^3\in\{0,1\}^3}\Big\{w(1,x_{1,1},\chi_1((1,0,0))) + w(0,x_{1,2},\chi_2((1,0,0))) + w(0,x_{1,3},\chi_3((1,0,0)))\}\Bigg) + (1-P_0(f_{\mathrm{S}})) \times \Big(\max_{x_1^3\in\{0,1\}^3\setminus\mathcal{X}_1^{100}}\{w(1,x_{1,1},\chi_1((1,0,0))) + w(0,x_{1,2},\chi_2((1,0,0))) + w(0,x_{1,3},\chi_3((1,0,0)))\}\Big)\Bigg], \tag{151}$$

$$\mathcal{X}_1^{100} = \arg\max_{x_1^3\in\{0,1\}^3}\{w(1,x_{1,1},\chi_1((1,0,0))) + w(0,x_{1,2},\chi_2((1,0,0))) + w(0,x_{1,3},\chi_3((1,0,0)))\}, \tag{152}$$

$$W_{111}(f_{\mathrm{S}},\chi_{\mathrm{S}}) = (1-p)^3\Bigg[P_0(f_{\mathrm{S}}) \times \Bigg(\max_{x_1^3\in\{0,1\}^3}\Big\{\sum_{t=1}^{3} w(1,x_{1,t},\chi_t((1,1,1)))\Big\}\Bigg) + (1-P_0(f_{\mathrm{S}})) \times \Bigg(\max_{x_1^3\in\{0,1\}^3\setminus\mathcal{X}_1^{111}}\Big\{\sum_{t=1}^{3} w(1,x_{1,t},\chi_t((1,1,1)))\Big\}\Bigg)\Bigg], \tag{153}$$

30

$$\mathcal{X}_1^{111} = \arg \max_{x_1^3 \in \{0,1\}^3} \{\sum_{t=1}^{3} w(1, x_{1,t}, \chi_t((1,1,1)))\}, \tag{154}$$

$$\begin{aligned} W_{011}(f_S \chi_S) = p(1-p)^2 \Bigg[& P_0(f_S) \times \Bigg(\max_{x_1^3 \in \{0,1\}^3} \Big\{ w(0, x_{1,1}, \chi_1((0,1,1))) + w(1, x_{1,2}, \chi_2((0,1,1))) \\ & + w(1, x_{1,3}, \chi_3((0,1,1))) \Big\} \Bigg) \\ & + (1 - P_0((f_S))) \times \Bigg(\max_{x_1^3 \in \{0,1\}^3 \setminus \mathcal{X}_1^{011}} \Big\{ w(0, x_{1,1}, \chi_1((0,1,1))) + w(1, x_{1,2}, \chi_2((0,1,1))) + \\ & w(1, x_{1,3}, \chi_3((0,1,1))) \Big\} \Bigg) \Bigg], \end{aligned} \tag{155}$$

$$\mathcal{X}_1^{011} = \arg \max_{x_1^3 \in \{0,1\}^3} \{w(0, x_{1,1}, \chi_1((0,1,1))) + w(1, x_{1,2}, \chi_2((0,1,1))) + w(1, x_{1,3}, \chi_3((0,1,1)))\}, \tag{156}$$

$$\begin{aligned} W_{101}(f_S, \chi_S) = p(1-p)^2 \Bigg[& P_0(f_S) \times \Bigg(\max_{x_1^3 \in \{0,1\}^3} \Big\{ w(1, x_{1,1}, \chi_1((1,0,1))) + w(0, x_{1,2}, \chi_2((1,0,1))) \\ & + w(1, x_{1,3}, \chi_3((1,0,1))) \Big\} \Bigg) \\ & + (1 - P_0(f_S)) \times \Bigg(\max_{x_1^3 \in \{0,1\}^3 \setminus \mathcal{X}_1^{101}} \Big\{ w(1, x_{1,1}, \chi_1((1,0,1))) + w(0, x_{1,2}, \chi_2((1,0,1))) + \\ & w(1, x_{1,3}, \chi_3((1,0,1))) \Big\} \Bigg) \Bigg], \end{aligned} \tag{157}$$

$$\mathcal{X}_1^{101} = \arg \max_{x_1^3 \in \{0,1\}^3} \Big\{ w(1, x_{1,1}, \chi_1((1,0,1))) + w(0, x_{1,2}, \chi_2((1,0,1))) + w(1, x_{1,3}, \chi_3((1,0,1))) \Big\}, \tag{158}$$

$$\begin{aligned} W_{110}(f_S, \chi_S) = p(1-p)^2 \Bigg[& P_0(f_S) \times \Bigg(\max_{x_1^3 \in \{0,1\}^3} \Big\{ w(1, x_{1,1}, \chi_1((1,1,0))) + w(1, x_{1,2}, \chi_2((1,1,0))) \\ & + w(0, x_{1,3}, \chi_3((1,1,0))) \Big\} \Bigg) \\ & + (1 - P_0(f_S)) \times \Bigg(\max_{x_1^3 \in \{0,1\}^3 \setminus \mathcal{X}_1^{110}} \Big\{ w(1, x_{1,1}, \chi_1((1,1,0))) + w(1, x_{1,2}, \chi_2((1,1,0))) \\ & + w(0, x_{1,3}, \chi_3((1,1,0))) \Big\} \Bigg) \Bigg], \end{aligned} \tag{159}$$

$$\mathcal{X}_1^{110} = \arg \max_{x_1^3 \in \{0,1\}^3} \{w(1, x_{1,1}, \chi_1((1,1,0))) + w(1, x_{1,2}, \chi_2((1,1,0))) + w(0, x_{1,3}, \chi_3((1,1,0)))\}. \tag{160}$$

In the case of Table V, $P_0(f_S)$ is given by

$$\begin{aligned} P_0(f_S) &= \mathrm{P}\big[x_0^3 = (0,0,0)\big] + \mathrm{P}\big[x_0^3 = (0,0,1)\big] + \mathrm{P}\big[x_0^3 = (0,1,0)\big] + \mathrm{P}\big[x_0^3 = (0,1,1)\big] \\ &\quad + \mathrm{P}\big[x_0^3 = (1,0,0)\big] + \mathrm{P}\big[x_0^3 = (1,0,1)\big] && (161) \\ &= p(2-p) && (162) \\ &\overset{p=\frac{1}{2}}{=} \frac{3}{4}. && (163) \end{aligned}$$

REFERENCES

[1] B. Larrousse and S. E. Lasaulce, "Coded power control: Performance analysis," in *Proc. of IEEE International Symposium on Information Theory*, Istanbul, Turkey, July 2013.

[2] B. Larrousse, A. Agrawal, and S. Lasaulce, "Implicit coordination in two-agent team problems; application to distributed power allocation," in *Proc. of 12th IEEE International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, 2014, pp. 579–584.

[3] T. Başar and S. Yüksel, *Stochastic Networked Control Systems*, Birhäuser, Ed. Springer, 2013, vol. XVIII.

[4] P. Cuff, H. H. Permuter, and T. M. Cover, "Coordination capacity." *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4181–4206, 2010.

[5] P. Cuff, "Distributed channel synthesis," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7071–7096, 2013.

[6] G. Kramer and S. A. Savari, "Communicating probability distributions," *IEEE Transactions on Information Theory*, vol. 53, no. 2, pp. 518–525, February 2007.

[7] T. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 752–772, May 1993.

[8] A. Bereyhi, M. Bahrami, M. Mirmohseni, and M. Aref, "Empirical coordination in a triangular multiterminal network," in *Proc. of IEEE International Symposium on Information Theory*, 2013, pp. 2149–2153.

[9] F. Haddadpour, M. H. Yassaee, A. Gohari, and M. R. Aref, "Coordination via a relay," in *Proc of IEEE International Symposium on Information Theory*, Boston, MA, July 2012, pp. 3048–3052.

[10] M. R. Bloch and J. Kliewer, "Strong coordination over a line network," in *Proc. IEEE International Symposium on Information Theory*, Istanbul, Turkey, July 2013, pp. 2319–2323.

[11] ——, "Strong coordination over a three-terminal relay network," accepted to *IEEE Information Theory Workshop*, May 2014.

[12] R. Blasco-Serrano, R. Thobaben, and M. Skoglund, "Polar codes for coordination in cascade networks," in *Proc. of International Zurich Seminar on Communications*, Zurich, Switzerland, March 2012, pp. 55–58.

[13] M. R. Bloch, L. Luzzi, and J. Kliewer, "Strong coordination with polar codes," in *Proc. of 50th Allerton Conference on Communication, Control, and Computing*, Monticello, IL, 2012, pp. 565–571.

[14] O. Gossner, P. Hernandez, and A. Neyman, "Optimal use of communication resources," *Econometrica*, vol. 74, no. 6, pp. 1603–1636, 2006.

[15] P. Cuff and L. Zhao, "Coordination using implicit communication," in *Proc. of IEEE Information Theory Workshop*, Paraty, Brazil, 2011, pp. 467–471.

[16] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1105–1115, Jun. 2002.

[17] S. Lasaulce and H. Tembine, *Game Theory and Learning for Wireless Networks : Fundamentals and Applications*. Academic Press, 2011.

[18] A. Zappone, S. Buzzi, and E. Jorswieck, "Energy-efficient power control and receiver design in relay-assisted DS/CDMA wireless networks via game theory," *IEEE Communications Letters*, vol. 15, no. 7, pp. 701–703, July 2011.

[19] G. Bacci, L. Sanguinetti, M. Luise, and H. Poor, "A game-theoretic approach for energy-efficient contention-based synchronization in OFDMA systems," *IEEE Transactions on Signal Processing*, vol. 61, no. 5, pp. 1258–1271, 2013.

[20] G. Scutari, D. Palomar, and S. Barbarossa, "The MIMO iterative waterfilling algorithm," *IEEE Transactions on Signal Processing*, vol. 57, no. 5, pp. 1917–1935, May 2009.

[21] P. Mertikopoulos, E. V. Belmega, A. L. Moustakas, and S. Lasaulce, "Distributed learning policies for power allocation in multiple access channels," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 96–106, January 2012.

[22] S. Gel'Fand and M. S. Pinsker, "Coding for channel with random parameters," *Probl. Contr. Inform. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[23] Y.-H. Kim, A. Sutivong, and T. Cover, "State amplification," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 1850–1859, 2008.

[24] C. Choudhuri, Y.-H. Kim, and U. Mitra, "Capacity-distortion trade-off in channels with state," in *Proc. of 48th Annual Allerton Conference on Communication, Control, and Computing*, 2010, pp. 1311–1318.

[25] ——, "Causal state amplification," in *Proc. of IEEE International Symposium on Information Theory Proceedings*, Saint Petersburg, Russia, August 2011, pp. 2110–2114.

[26] C. Choudhuri and U. Mitra, "Action dependent strictly causal state communication," in *Proc. of IEEE International Symposium on Information Theory*, 2012, pp. 3058–3062.

[27] P. Cuff and C. Schieler, "Hybrid codes needed for coordination over the point-to-point channel," in *Proc. of 49th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, 2011, pp. 235–239.

[28] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 2006.

[29] M. L. Treust and S. Lasaulce, "A repeated game formulation of energy-efficient decentralized power control," *IEEE Transactions on Wireless Communications*, vol. 9, no. 9, pp. 2860–2869, September 2010.

[30] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.

[31] A. El Gamal and Y. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[32] A. S. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1639–1667, June 2002.

[33] D. Goodman and N. Mandayam, "Power control for wireless data," *IEEE Personal Communications*, vol. 7, no. 2, pp. 45–54, April 2000.

[34] B. Fourestié, "Communication switching method, access point, network controller and associated computer programs," France Telecom, Patent WO 2008/081137, Dec. 2007.

[35] N. Merhav and S. Shamai, "On joint source-channel coding for the Wyner-Ziv source and the Gel'fand-Pinsker channel," *IEEE Transactions on Information Theory*, vol. 49, no. 11, pp. 2844–2855, 2003.

[36] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, 2005.

[37] Q. Li, D. Gesbert, and N. Gresset, "Joint precoding over a master-slave coordination link," in *Proc. of the IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Florence, Italy, May 2014.

[38] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.

[39] T. M. Cover and M. Chiang, "Duality between channel capacity and rate distortion with two-sided state information," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1629–1638, 2002.

[40] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[41] A. Gjendemsj, D. Gesbert, G. E. Oien, and S. G. Kiani, "Binary power control for sum rate maximization over multiple interfering links," *IEEE Transactions on Wireless Communications*, vol. 7, no. 8, pp. 3164–3173, 2008.

[42] F. Meshkati, H. Poor, and S. Schwartz, "Energy-efficient resource allocation in wireless networks: An overview of game theoretic approaches," *IEEE Signal Processing Magazine*, vol. 58, pp. 58–68, May 2007.

[43] E. V. Belmega and S. Lasaulce, "Energy-efficient precoding for multiple-antenna terminals," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 329–340, January 2011.

E.2 Conférences internationales avec comité de lecture et actes

E.2.1 IEEE-ISIT 2013

2013 IEEE International Symposium on Information Theory

Coded Power Control: Performance Analysis

Benjamin Larrousse and Samson Lasaulce
L2S (CNRS – Supélec – Univ. Paris Sud 11), 91191 Gif-sur-Yvette, France
Email: {larrousse, lasaulce}@lss.supelec.fr

***Abstract*—In this paper, we introduce the general concept of coded power control (CPC) in a particular setting of the interference channel. Roughly, the idea of CPC consists in embedding information (about the channel state) into the transmit power levels themselves: in this new framework, provided the power levels of a given transmitter can be observed by other transmitters, a sequence of power levels of the former can therefore be used to coordinate the latter. To assess the limiting performance of CPC (and therefore the potential performance brought by this new approach), we derive, as a first step towards many extensions of the present work, a general result which not only concerns power control (PC) but also any scenario involving two decision-makers (DMs) which communicate through their actions and have the following information and decision structures. We assume that the DMs want to maximize the average of an arbitrarily chosen instantaneous payoff function which depends on the DMs' actions and the state realization. DM 1 is assumed to know the state non-causally (e.g., the channel state) which affects the common payoff while DM 2 has only a strictly causal knowledge of it. DM 1 can only use its own actions (e.g., power levels) to inform DM 2 about its best action in terms of payoff. Importantly, DM 2 can only monitor the actions of DM 1 imperfectly and DM 2 does not need to be observed by DM 1. The latter assumption leads us to exploiting Shannon-theoretic tools in order to generalize an existing theorem which provides the information constraint under which the average payoff is maximized. The derived result is then exploited to fully characterize the performance of good CPC policies for a given instance of the interference channel.**

I. Introduction

Consider two decision-makers (DMs) and that each of them has to select actions or take decisions repeatedly to reach a common objective namely, to maximize an average payoff function. Furthermore, assume that there might be an interest for them in exchanging information e.g., about the future events which can affect their payoff but that no dedicated communication channel is available for this purpose. Therefore, the only way to communicate for a DM is to use his own actions. Although the idea of communicating through actions seems to be quite natural and is in fact used more or less implicitly in real life scenarios e.g., in economics (see [1]), it appears that, apart from a few exceptions focused on specific problems of control (see e.g., [2]), it has obviously not penetrated yet engineering problems and definitely not wireless communications. It turns out that important wireless problems such as power control (PC) or radio resource allocation can draw much benefits from being revisited from the new perspective of communication through actions. Because of its importance and ability to easily illustrate the proposed approach, the problem of PC in interference networks has been selected for the application of the main and general result derived in this paper. Note that the latter concerns any decision-making problem which has the same structure (see Sec. II) and generalizes [3] (and [4] which exploits the same theorem); in the (game-theoretical) setting of [3], DM 1 perfectly monitors the actions of DM 2 and conversely.

In the context of PC, the DMs are transmitters (Txs) and the system state is typically given by the state of the communication channel between the Txs and receivers (Rxs); we will use the term DM (resp. Tx) when the general case (resp. the specific case of PC) is concerned. Quite often, each Tx possesses a partial knowledge of the channel state and, in general, there is an incentive for the Txs to exchange the corresponding knowledge between them. Coded power control[1] (CPC) assumes that this knowledge is transferred from one Tx to another (or others) by encoding the information of the former into a sequence of power levels which are observed by the latter. CPC is, in particular, relevant in cognitive radio (CR) settings. In typical CR scenarios, the primary Tx is assumed to be passive and the secondary Tx adapts to what it observes. But, it might be of interest to design primary Txs which coordinate in an active manner the usage of radio resources, which is exactly what CPC allows; one of the salient features of CPC is that interference can be managed directly in the radio-frequency domain and does not require baseband detection or decoding, which is very useful in heterogeneous networks. Another body of works which can be mentioned is given by works on distributed PC and especially those on best response dynamics (BRD) algorithms which include the original iterative water-filling algorithm [6]. Existing BRD algorithms implementations for PC (see e.g., [7][8][9]) typically assume SINR (signal-to-noise plus interference ratio) feedback and individual channel state information (CSI) and do not exploit the key idea of communicating through the power levels. Encoding power levels allows one to construct PC policies possessing at least three salient features which are generally not available for BRD-based PC: there is no convergence problem and this whatever the payoff functions; efficient solutions (e.g., in terms of sum-payoff) can be obtained; both the cases of discrete and continuous power levels can be easily treated. Since we focus on optimal PC policies and make the choice of an asymmetric information structure whereas BRD algorithms rely on a symmetric one, no explicit comparison with BRD algorithms is conducted but CPC can be applied to symmetric scenarios as well.

[1]The first document to disclose this idea is [5].

978-1-4799-0446-4/13/$31.00 ©2013 IEEE

2013 IEEE International Symposium on Information Theory

II. Problem statement

Consider two DMs which have to select actions repeatedly (over T stages) and want to coordinate through their actions. Let $\mathcal{X}_j$, $|\mathcal{X}_j| < \infty$, the action alphabet of DM $j \in \{1,2\}$, and $\mathcal{X}_0$, $|\mathcal{X}_0| < \infty$, the random state alphabet. The states are assumed to be i.i.d. and generated from a random variable X_0 whose realizations are in $\mathcal{X}_0$ and distribution is denoted by ρ. Note that the finiteness assumption is not only realistic (e.g., power levels are discrete in modern cellular systems) but also allows the continuous case to be treated by using classical arguments [10]. The strategies of DM 1 and 2 are sequences of mappings, $(\sigma_i, \tau_i)_{i\geq 1}$, which are respectively defined by:

$$\left\{ \begin{array}{rcccc} \sigma_i & : & \mathcal{X}_0^T \times \mathcal{X}_1^{i-1} & \rightarrow & \mathcal{X}_1 \\ \tau_i & : & \mathcal{X}_0^{i-1} \times \mathcal{Y}^{i-1} \times \mathcal{X}_2^{i-1} & \rightarrow & \mathcal{X}_2 \end{array} \right. \quad (1)$$

where T is the total number of stages, $i \in \{1, ..., T\}$ is the stage index, and $\mathcal{Y}$, $|\mathcal{Y}| < \infty$, is the observation alphabet of DM 2. The definition of the strategy for DM 1 indicates that we assume a non-causal knowledge of the state. The most typical situation in PC is to assume that two phases are available (training phase, action phase) and one state is known in advance to adjust the power level. This special case can be obtained by setting $T = 2$ that is, $i \in \{1, 2\}$. There are many reasons why we consider here that T might be greater than two. We will only provide three of them, which better explains how the non-causality assumption may be understood. First, the result derived in Sec. III can be used for a large variety of settings and not only PC. Second, the proposed approach can be applied to the case where the state is not i.i.d. (e.g., to the $B-$stage block i.i.d. case, $B \geq 1$). Indeed, there exist wireless communication standards which assume the channel to be constant over several time-slots and the proposed approach suggests that gains can be obtained by varying the power level from time-slot to time-slot even if the channel is constant. Third, it becomes more and more common to exploit the forecasted trajectory of a mobile user to optimize the system [11], which makes our approach relevant when the channel state is interpreted as the path loss. Concerning the chosen definition for the strategy of DM 2, several comments are in order. First, note that DM 2 is not assumed to monitor actions of DM 1 perfectly. Rather, they are monitored through an observation channel which is assumed to be discrete, memoryless, and to verify $P(y|x_0, x_1, x_2) = \Gamma(y|x_1)$, where $y \in \mathcal{Y}$ is a realization of the channel output associated with the input (x_0, x_1, x_2). Second, note that, the strategy of DM 2 is defined such that it can choose an action at every stage and not only at the end of a block or sequence of stages as it would be the case for a classical block decoder. Therefore, contrarily to [3], DM 1 does not need to observe the actions of DM 2 and DM 2 has only access to imperfect observations of the actions chosen by DM 1. Interestingly, we will see that the fact that DM 1 does not observe DM 2 induces no performance loss in terms of payoff.

The instantaneous or stage payoff function for the DMs is denoted by $w(x_0, x_1, x_2)$. Since the state is random, we will consider as general case the problem of reaching a certain performance level in terms of expected payoff $\mathbb{E}[w] = \sum_{(x_0,x_1,x_2)} P(x_0, x_1, x_2)w(x_0, x_1, x_2)$. Roughly, the task of DM 1 is to maximize the expected payoff by finding the best tradeoff between reaching a good payoff for the current stage and revealing enough information about the future realizations of the state to coordinate for the next stages. The ability for two DMs to coordinate their actions i.e., to reach a certain value for the expected payoff can be translated in terms of joint distribution over $\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2$, which leads us to the notion of implementable distribution [3].

Definition 1 (Implementability)**.** *The distribution $\overline{Q}(x_0, x_1, x_2)$ is implementable if there exists a pair of strategies $(\sigma_i, \tau_i)_{i\geq 1}$ such that as $t \rightarrow \infty$ we have for all (x_0, x_1, x_2),*

$$\frac{1}{t}\sum_{i=1}^{t}\sum_{y\in\mathcal{Y}} P_{X_{0,i},X_{1,i},X_{2,i},Y_i}(x_0, x_1, x_2, y) \rightarrow \overline{Q}(x_0, x_1, x_2) \quad (2)$$

where $P_{X_{0,i},X_{1,i},X_{2,i},Y_i}$ is the joint distribution induced by $(\sigma_i, \tau_i)_{i\geq 1}$ at stage i.

Importantly, note that the set of feasible payoffs which are reachable asymptotically $\lim_{T\rightarrow\infty} \frac{1}{T}\sum_i w(x_0(i), x_1(i), x_2(i))$ is a linear image of the set of implementable distributions under the expectation operator. Therefore, a certain value, say $\overline{w}$, is reachable asymptotically if and only if there exists an implementable distribution Q such that $\overline{w} = \mathbb{E}_Q[w]$. The goal of the next section is precisely to characterize the set of reachable expected payoffs $\mathbb{E}_Q[w] = \sum_{(x_0,x_2,x_2,y)} \overline{Q}(x_0, x_1, x_2)\Gamma(y|x_1)w(x_0, x_1, x_2)$, which thus amounts to characterizing the set of implementable distributions over $\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2$.

III. Main analytical result

Notation: $\Delta(\mathcal{A})$ will stand for the set of distributions over the generic discrete set $\mathcal{A}$. Using this notation, the main analytical result of this paper states as follows.

Theorem 1. *Let $\overline{Q} \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ with $\sum_{(x_1,x_2)} \overline{Q}(x_0, x_1, x_2) = \rho(x_0)$. The distribution $\overline{Q}$ is implementable if and only if there exists $Q \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})$ which verifies the following information constraint:*

$$I_Q(X_0; X_2) \leq I_Q(X_1; Y|X_0, X_2) \quad (3)$$

where the arguments of the mutual information $I_Q(.)$ are defined from Q and $Q(x_0, x_1, x_2, y) = \overline{Q}(x_0, x_1, x_2)\Gamma(y|x_1)$.

A. Proof of Theorem 1

Converse proof: We first start with providing a lemma which is used at the end of the converse and concludes the section.

Lemma 1. *The function $\Phi : Q \mapsto I_Q(X_0; X_2) - I_Q(X_1; Y|X_0, X_2)$ is convex over the set of distributions $Q \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})$ that verify $\sum_{(x_1,x_2,y)} Q(x_0, x_1, x_2, y) =$*

2013 IEEE International Symposium on Information Theory

$\rho(x_0)$ and $Q(x_0,x_1,x_2,y) = \Gamma(y|x_1)P(x_0,x_1,x_2)$, with ρ and Γ fixed.

Proof of Lemma 1: The function Φ can be rewritten as $\Phi(Q) = H_Q(X_0) - H_Q(Y,X_0|X_2) + H_Q(Y|X_0,X_2,X_1)$. The first term $H_Q(X_0) = -\sum_{x_0}\rho(x_0)\log\rho(x_0)$ is a constant w.r.t. Q. The third term is linear w.r.t. Q since, with Γ fixed,

$$\begin{aligned} H_Q(Y|X_0,X_2,X_1) &= \\ &- \sum_{x_0,x_1,x_2,y} Q(x_0,x_1,x_2,y)\log P(y|x_0,x_1,x_2) \\ &= - \sum_{x_0,x_1,x_2,y} Q(x_0,x_1,x_2,y)\log\Gamma(y|x_1) \quad (4) \end{aligned}$$

It is therefore sufficient to prove that $H_Q(Y,X_0|X_2)$ is concave. Let $\lambda_1 \in [0,1]$, $\lambda_2 = 1-\lambda_1$, $(Q_1,Q_2) \in \Delta^2(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y})$ and $Q = \lambda_1 Q_1 + \lambda_2 Q_2$. By using the standard notation $A^0 = \emptyset$, $A^n = (A_1,...,A_n)$, we have that:

$$\begin{aligned} &H_Q(Y,X_0|X_2) = -\sum_{x_0,x_2,y}\left(\sum_{x_1,i}\lambda_i Q_i(x_0,x_1,x_2,y)\right) \\ &\log\left[\frac{\sum_{x_1,i}\lambda_i Q_i(x_0,x_1,x_2,y)}{\sum_i \lambda_i P_{X_2}^{Q_i}(x_2)}\right] \quad (5) \\ &= -\sum_{x_0,x_2,y}\left(\sum_i \lambda_i \sum_{x_1} Q_i(x_0,x_1,x_2,y)\right) \\ &\log\left[\frac{\sum_i \lambda_i \sum_{x_1} Q_i(x_0,x_1,x_2,y)}{\sum_i \lambda_i P_{X_2}^{Q_i}(x_2)}\right] \quad (6) \\ &\geq -\sum_i \lambda_i \sum_{x_0,x_2,y}\left(\sum_{x_1} Q_i(x_0,x_1,x_2,y)\right) \\ &\log\left[\frac{\lambda_i \sum_{x_1} Q_i(x_0,x_1,x_2,y)}{\lambda_i P_{X_2}^{Q_i}(x_2)}\right] \quad (7) \\ &= -\sum_i \lambda_i \sum_{x_0,x_2,y}\left(\sum_{x_1} Q_i(x_0,x_1,x_2,y)\right) \\ &\log\left[\frac{\sum_{x_1} Q_i(x_0,x_1,x_2,y)}{P_{X_2}^{Q_i}(x_2)}\right] \quad (8) \\ &= \lambda_1 H_{Q_1}(Y,X_0|X_2) + \lambda_2 H_{Q_2}(Y,X_0|X_2) \quad (9) \end{aligned}$$

where the inequality comes from the log sum inequality [10]. ∎

Now we want to prove that if $\overline{Q}$ is implementable, then Q has to verify the information constraint. Assuming $\overline{Q}$ is implementable means that there exists $(\sigma_i,\tau_i)_{i\geq 1}$ such that the empirical distribution $P^{(t)}_{X_0,X_1,X_2,Y}(.) = \frac{1}{t}\sum_{i=1}^t P_{X_{0,i},X_{1,i},X_{2,i},Y_i}(.)$ can be made arbitrarily close to Q; this argument is used at the end of the proof. We have:

$$\begin{aligned} &\sum_{i=1}^t I_{P_{X_{0,i},X_{1,i},X_{2,i},Y_i}}(X_0;X_2) = \sum_{i=1}^t I(X_{0,i};X_{2,i}) \quad (10) \\ &\overset{(a)}{=} H(X_0^t) - \sum_{i=1}^t H(X_{0,i}|X_{2,i}) \quad (11) \\ &= H(X_0^t,Y^t,X_2^t) - H(Y^t,X_2^t|X_0^t) - \sum_{i=1}^t H(X_{0,i}|X_{2,i}) \quad (12) \\ &\leq H(X_0^t,Y^t,X_2^t) - H(Y^t|X_0^t) - \sum_{i=1}^t H(X_{0,i}|X_{2,i}) \quad (13) \\ &\leq H(X_0^t,Y^t,X_2^t) - H(Y^t|X_0^t,X_1^t,X_2^t) - \sum_{i=1}^t H(X_{0,i}|X_{2,i}) \quad (14) \\ &\overset{(b)}{=} H(X_0^t,Y^t,X_2^t) - \sum_{i=1}^t H(X_{0,i}|X_{2,i}) \\ &\quad - \sum_{i=1}^t H(Y_i|X_{0,i},X_{1,i},X_{2,i}) \quad (15) \\ &\overset{(c)}{\leq} \sum_{i=1}^t H(X_{0,i},Y_i,X_{2,i}|X_{2,i}) - H(X_{0,i}|X_{2,i}) \\ &\quad - H(Y_i|X_{1,i},X_{0,i},X_{2,i}) \quad (16) \\ &= \sum_{i=1}^t H(X_{0,i},Y_i|X_{2,i}) - H(X_{0,i}|X_{2,i}) \\ &\quad - H(Y_i|X_{1,i},X_{0,i},X_{2,i}) \quad (17) \\ &= \sum_{i=1}^t I(X_{1,i};Y_i|X_{0,i},X_{2,i}) \quad (18) \\ &= \sum_{i=1}^t I_{P_{X_{0,i},X_{1,i},X_{2,i},Y_i}}(X_1;Y|X_0,X_2) \quad (19) \end{aligned}$$

where: (a) comes from the fact that $(X_{0,i})_i$ is i.i.d. and the chain rule for entropy; (b) holds because the observation channel from DM 1 to DM 2 is assumed to be discrete and memoryless namely, $P(y^t|x_0^t,x_1^t,x_2^t) = \prod_{i=1}^t p(y_i|x_{0,i},x_{1,i},x_{2,i})$; (c) holds by the chain rule and because $X_{2,i}$ is a deterministic function of the past: $X_{2,i} = \tau_i(X_{0,1},Y_1,X_{2,1},\ldots,X_{0,i-1},Y_{i-1},X_{2,i-1})$. Now, since Φ is convex (by Lemma 1), we know that

$$\begin{aligned} I_{P^{(t)}_{X_0,X_1,X_2,Y}}(X_1;Y|X_0,X_2) - I_{P^{(t)}_{X_0,X_1,X_2,Y}}(X_0;X_2) \geq \\ \frac{1}{t}\sum_{i=1}^t I_{P_{X_{0,i},X_{1,i},X_{2,i},Y_i}}(X_1;Y|X_0,X_2) \\ - I_{P_{X_{0,i},X_{1,i},X_{2,i},Y_i}}(X_0;X_2) \quad (20) \end{aligned}$$

The converse follows by observing that the first term of the above inequality can be made arbitrarily close to $I_Q(X_1;Y|X_0,X_2) - I_Q(X_0;X_2)$ and the second term has been proven to be non-negative. ∎

Implementability (sketch): The goal here is to prove that if the information constraint is verified, then an implementable pair of strategies $(\sigma_i,\tau_i)_{i\geq 1}$ can be found. Therefore, in

2013 IEEE International Symposium on Information Theory

contrast with the converse, finding a particular code such as a block code with long codewords is sufficient, which allows one to reuse the standard machinery for the transmission of distorted sources. The methodology is therefore to assume $T = nB$ is large where n is the codeword length and B the number of blocks, and exploit the block Markov coding idea as a particular case of general strategy. By denoting b as the block index, the strategy of DM 2 is taken to be the mapping $\psi : (x_0^n(b), y^n(b), x_2^n(b)) \mapsto x_2^n(b+1)$. Since DM 1 knows the state for the next block whose index is $b+1$ (and also $x_0^n(b)$), the strategy of DM 1 is taken to be the mapping $\phi : (x_0^n(b+1), x_0^n(b), \widehat{x}_2^n(b)) \mapsto x_1^n(b)$ where $\widehat{x}_2^n(b) = \psi(x_0^n(b-1), x_1^n(b-1), \widehat{x}_2^n(b-1))$. At last, for block $b = 0$, DM 2 chooses an arbitrary codeword which is known to DM 1s. Under this setting, it can be shown that if Q meets the information constraint then the empirical distribution $Q^{(T)}$ which is induced by a separate source channel coding i.e.,

$$Q^{(T)}(v) = \frac{1}{nB}\big[\mathcal{N}\big(v \mid x_0^n(0), x_1^n(0), x_2^n(0), y^n(0)\big) + \sum_{b=1}^{B-1} \mathcal{N}\big(v \mid x_0^n(b), x_1^n(b), x_2^n(b), y^n(b)\big)\big]$$

converges to Q, where $\mathcal{N}(v|v^n)$ is a notation for counting the occurrences of v in v^n, $v = (x_0, x_1, x_2, y)$ here. ∎

B. Comments on Theorem 1

Theorem 1 can be interpreted as follows. DM 2's actions (represented by X_2) correspond to a joint source-channel decoding operation with distortion on the information source (which is represented by X_0). To be reachable, the distortion rate has to be less than the transmission rate allowed by the channel whose input and output are respectively represented by X_1 and Y. Therefore, the pair $S = (X_0, X_2)$ seems to play the same role as the side information in channels with state. Indeed, the implementability proof shows that DM 1 uses in particular $(x_0^n(b), \widehat{x}_2^n(b))$ while DM 2 uses $(x_0^n(b), x_2^n(b))$. Asymptotically, the encoder (DM 1) and decoder (DM 2) have the same side information; this observation is one of the key elements of the implementability proof. It explains why the fact that DM 1 does not need to observe DM 2 does not induce any performance loss. Furthermore, note that $x_0^n(b+1)$, which plays the role of the message to be encoded, is independent of the side information. Classical coding schemes (such as block Markov coding) can thus be re-exploited. However, the above arguments fails for the converse proof which has to deal with arbitrary coding schemes or strategies. It can no longer be assumed that the side information be independent of the information source vector. This is one of the reasons why the converse proof has to be rethought. Another reason is that classical results (such as Fano's inequality) which rely on block decoding are not exploitable anymore since DM 2 has to be able to act (to decode) at any stage or time instance.

As another type of comments on Theorem 1, it can be noted that the information constraint has a very attractive property: the problem of maximizing the expected payoff takes a particularly simple form. Indeed, by defining a one-to-one mapping between the quadruplets (x_0, x_1, x_2, y) and the finite set $\{1, 2, ..., L\}$, $L = |\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}|$, the optimization problem of interest can be described as follows:

$$\begin{array}{lrcl} \text{minimize} & -\mathbb{E}_{\underline{q}}[w] = -\sum_{\ell=1}^{L} q_\ell w_\ell & & \\ \text{subject to} & I_{\underline{q}}(X_0; X_2) - I_{\underline{q}}(X_1; Y|X_0, X_2) & \leq & 0 \\ & -q_\ell & \leq & 0 \\ & -1 + \sum_{\ell=1}^{L} q_\ell & = & 0 \\ & \forall x_0,\ \sum_{\ell \in \mathcal{L}_{X_0}(x_0)} q_\ell - \rho(x_0) & = & 0 \\ & \forall (x_1, y),\ \frac{\sum_{\ell \in \mathcal{L}_{X_1,Y}(x_1,y)} q_\ell}{\sum_{\ell \in \mathcal{L}_{X_1}(x_1)} q_\ell} - \Gamma(y|x_1) & = & 0 \end{array} \tag{21}$$

where q_ℓ is the probability of a given quadruplet (x_0, x_1, x_2, y), w_ℓ is the value of the corresponding payoff, the vector $\underline{q} = (q_1, ..., q_L)$ represents the distribution Q, and the sets of indices $\mathcal{L}_{X_0}(x_0)$, $\mathcal{L}_{X_1,Y}(x_1, y)$, $\mathcal{L}_{X_1}(x_1)$ merely translate the marginalization conditions. By Lemma 1, it follows that the above optimization problem is convex, which makes easy the determination of the information-constrained maximum of the expected payoff. A simple and useful upper bound for this maximum is $\mathbb{E}_\rho \max_{(x_1,x_2)} w(x_0, x_1, x_2)$. This bound will be referred to as the costless communication case in Sec. IV. Indeed, this bound can be attained in the ideal scenario where: given the knowledge of the coming state x_0, DM 1 computes an optimal solution for the action pair for the coming stage $(x_1^*, x_2^*) \in \arg\max_{(x_1,x_2)} w(x_0, x_1, x_2)$ and can inform DM 2 of x_2^* without any cost. If the state is stationary for say S stages and $\mathcal{X}_1 = \mathcal{X}_2$, a simple strategy for DM 1 can be as follows: $x_1(1) = x_2^*$, $x_1(2) = x_1^*$, ..., $x_1(S) = x_1^*$. This allows DM 2 to choose an optimal action for $i \in \{2, ..., S\}$. It can be shown that considering the $S-$stage block i.i.d. case amounts to multiplying the left term of (3) by $\frac{1}{S}$, which makes the information constraint arbitrarily mild as S grows large.

IV. Application to power control over interference channels

The main goal is to assess the performance of simple CPC policies and those of good policies; the performance of the latter is obtained by exploiting Theorem 1. A flat-fading interference channel (IC) with two Tx-Rx pairs is considered. Transmissions are assumed to be time-slotted and synchronized. For $j \in \{1, 2\}$ and "$k = -j$" ($-j$ stands for the Tx other than j), the SINR at receiver j at a given stage writes as $\text{SINR}_j = \frac{g_{jj} x_j}{N + g_{kj} x_k}$ where $x_j \in \mathcal{X}_j^{\text{IC}} = \{0, P_{\max}\}$ is the power level chosen by Tx j, g_{jk} represents the channel gain of link jk, and N the noise variance. We assume that: $g_{jk} \in \{g_{\min}, g_{\max}\}$ is i.i.d. and Bernouilli distributed $g_{jk} \sim \mathcal{B}(p_{jk})$ with $P(g_{jk} = g_{\min}) = p_{jk}$. We define SNR[dB] $= 10 \log_{10} \frac{P_{\max}}{N}$ and set $g_{\min} = 0.1$, $g_{\max} = 1.9$, $N = 1$. The low and high interference regimes (LIR, HIR) are respectively defined by $(p_{11}, p_{12}, p_{21}, p_{22}) = (0.5, 0.9, 0.9, 0.5)$ and

2013 IEEE International Symposium on Information Theory

$(p_{11}, p_{12}, p_{21}, p_{22}) = (0.5, 0.1, 0.1, 0.5)$. The assumed payoff is $w^{\text{IC}}(x_0, x_1, x_2) = \sum_{j=1}^{2} f(\text{SINR}_j(x_0, x_1, x_2))$ where $f(a) = \log(1+a)$ unless stated otherwise. At last, we assume that $Y \equiv X_1$. We consider four CPC policies :

► the full power control (FPC) policy $x_j = P_{\max}$ for every stage. FPC requires no CSI at all;

► the semi-coordinated PC (SPC) policy $x_2 = P_{\max}$, $x_1^\dagger \in \arg\max_{x_1} w^{\text{IC}}(x_0, x_1, P_{\max})$. SPC requires the knowledge of the current state realization at Tx1;

► the optimal CPC policy (OCPC) whose performance is obtained, in particular, when the problem has the information structure of Theorem 1;

► the costless communication case (see Sec. III-B) for which the maximum of w^{IC} can be reached at any stage. Fig. 1 and 2 depict the relative gain in % in terms of average payoff versus SNR[dB] which is obtained by FPC, SPC, OCPC, and costless case. Compared to FPC, gains are very significant whatever the interference regime and provided the SNR has realistic values. Compared to SPC, the gain is of course less impressive since SPC is precisely a coordinated PC scheme but, in the HIR and when the communication cost is negligible, gains as high as 25% can be obtained with $f(a) = \log(1+a)$ and 45% with $f(a) = a$.

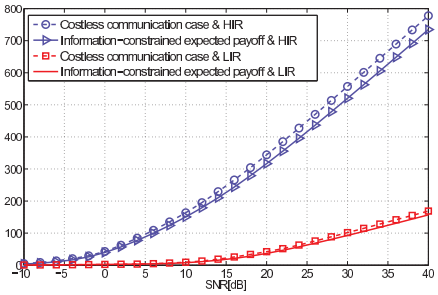

Fig. 1. Relative gain in terms of average payoff ("OCPC/FPC − 1" in [%]) vs SNR[dB] obtained with CPC (with and without communication cost) when the reference power control policy is to transmit at full power (FPC).

V. Concluding Remarks

Although some assumptions made in this paper might be too restrictive in some application scenarios, it is essential to understand that the used methodology to derive the optimal performance is general. It can be applied to analyze the performance of coded power allocation, coded interference alignment, etc, with other information structures and by considering $N \geq 2$ individual payoffs instead of a common one (e.g., in a game-theoretic setting [3]). The methodology to assess the performance of good coded policies consists in deriving the right information constraint(s) by building the proof on Shannon theory for the problem of multi-source coding with distortion over multi-user channels wide side information and then to use this constraint to find an information-constrained

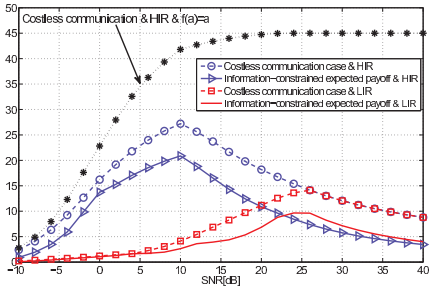

Fig. 2. The difference with Fig. 1 is that the reference power control policy is the SPC policy. Additionally, the top curve is obtained with $f(a) = a$.

maximum of the payoff (common payoff case) or the set of Nash equilibrium points which are compatible with the constraint (non-cooperative game case). Note that assuming the state(s) to be i.i.d. from stage to stage leads in fact to the worst-case scenario for the information constraint. On the other hand, the costless communication case provides an upper bound for the expected payoff. As a key observation of this paper, the communication structure of a multi-person decision-making problem corresponds in fact to a multiuser channel. Therefore, multi-terminal Shannon theory is not only relevant for pure communication problems but also for any multi-person decision-making problem. The above observation also opens new challenges for Shannon-theorists since decision-making problems define new communication channels.

Acknowledgment

The authors would like to thank Prof. Olivier Gossner for providing interesting feedbacks on his work [3].

References

[1] C. A. Sims, "Implications of rational inattention," *Journal of Monetary Economics*, vol. 50, no. 3, pp. 665–690, 2003.

[2] P. Grover, "Actions can speak more clearly than words," Ph.D. dissertation, Univ. of California at Berkeley, Jan. 2011.

[3] O. Gossner, P. Hernandez, and A. Neyman, "Optimal use of communication resources," *Econometrica*, vol. 74, no. 6, pp. 1603–1636, 2006.

[4] P. Cuff and L. Zhao, "Coordination using implicit communication," in *IEEE Information Theory Workshop*, Oct. 2011.

[5] S. M. Perlaza, S. Lasaulce, and G. Salingue, "Strategic communication in wireless self-configuring networks," Orange, Patent 54719, Nov. 2010.

[6] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1105–1115, June 2002.

[7] S. Lasaulce and H. Tembine, *Game Theory and Learning for Wireless Networks : Fundamentals and Applications*. Academic Press, 2011.

[8] A. Zappone, S. Buzzi, and E. Jorswieck, "Energy-efficient power control and receiver design in relay-assisted DS/CDMA wireless networks via game theory," *IEEE Comm. Lett.*, vol. 15, no. 7, pp. 701–703, July 2011.

[9] G. Bacci, L. Sanguinetti, M. Luise, and H. V. Poor, "A game-theoretic approach for energy-efficient contention-based synchronization in OFDMA systems," *IEEE Trans. on Signal Processing*, 2013.

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 2006.

[11] B. Fourestié, "Communication switching method, access point, network controller and associated computer programs," France Telecom, Patent WO 2008/081137, Dec. 2007.

E.2.2 IEEE-WiOpt 2014

WNC3 2014: International Workshop on Wireless Networks: Communication, Cooperation and Competition

Implicit Coordination in Two-Agent Team Problems; Application to Distributed Power Allocation

Benjamin Larrousse, Achal Agrawal, and Samson Lasaulce
L2S (CNRS – Supélec – Univ. Paris Sud 11), 91192 Gif-sur-Yvette, France
Email: {larrousse, agrawal, lasaulce}@lss.supelec.fr

***Abstract*—The central result of this paper is the analysis of an optimization problem which allows one to assess the limiting performance of a team of two agents who coordinate their actions. One agent is fully informed about the past and future realizations of a random state which affects the common payoff of the agents whereas the other agent has no knowledge about the state. The informed agent can exchange his knowledge with the other agent only through his actions. This result is applied to the problem of distributed power allocation in a two-transmitter $M-$band interference channel, $M \geq 1$, in which the transmitters (who are the agents) want to maximize the sum-rate under the single-user decoding assumption at the two receivers; in such a new setting, the random state is given by the global channel state and the sequence of power vectors used by the informed transmitter is a code which conveys information about the channel to the other transmitter.**

I. Introduction and case study of interest

Consider two agents or decision-makers who interact over a time period composed of a large number of stages or time-slots. At each stage $t \in \mathbb{N}$, agent $i \in \{1,2\}$ chooses an action $x_i \in \mathcal{X}_i$, $|\mathcal{X}_i| < +\infty$. The resulting (instantaneous) agents' common payoff is $u(x_0, x_1, x_2)$ where $x_0 \in \mathcal{X}_0$, $|\mathcal{X}_0| < +\infty$, is the realization of a random state for the considered stage, and w is a real valued function. This random state is assumed to be an i.i.d. random process. To assess the (theoretical) limiting achievable coordination performance, it is assumed that one agent, agent 1, knows beforehand and perfectly all the realizations of the random state. Note that, in practice, the sole knowledge of the next realization of the random state is already very useful, just as in conventional power control problems. On the other hand, agent 2 does not know the state at all and can only be informed about it by observing the actions of agent 1. The performance analysis of this problem leads to deriving an information-theoretic constraint. The case of perfect observation is treated in [1] while the generalization to noisy observations is conducted in [2]; to be precise, both references assume that agent 2 has a strictly causal knowledge of the state but it can be shown that not having any knowledge about the state's realizations at all induces no limiting performance loss [3]. Reference [2] also states an optimization problem which essentially amounts to maximizing the long term payoff function under some constraints but this optimization problem is not analyzed. One of the purposes of this paper is precisely to study this general problem in detail. This will allow one to specialize it for the specific problem of power allocation in an important setting of cognitive radio.

The application of interest in this paper corresponds to a scenario which involves two transmitter-receiver pairs whose communications interfere each other. The communication system under consideration is modeled by an $M-$band interference channel, $M \geq 1$, as depicted in Fig. 1. In contrast with the vast majority of related works on distributed power allocation over multi-band channels (starting with the pioneering work [4]), the set of power allocation vectors at a transmitter is assumed to be discrete and finite (namely, $|\mathcal{X}_i| < +\infty$) instead of being continuous. This choice is motivated by many applications (see e.g., [5][6][7][8]) and by well-known results in information theory [9] which show that the continuous case generally follows from the discrete case by calling quantization arguments. We also assume that channel gains, as defined by Fig. 1, lie in discrete sets; this is also well motivated by practical applications such as cellular systems in which quantities such as the channel quality indicator are used. Therefore, for the considered case study, x_0 is given by the vector of all channel gains g_{ij}^m, $(i,j) \in \{1,2\}^2, m \in \{1,2,...,M\}$, and lies in a finite discrete set (denoted by $\mathcal{X}_0$).

The paper is organized as follows. In Sec. II, we introduce and solve the general optimization problem of interest. In Sec. III, we apply the general result of Sec. II to a special case of payoff function and action sets for the agents. This special case corresponds to the problem of power allocation in a cognitive radio scenario. Sec. IV concludes the paper.

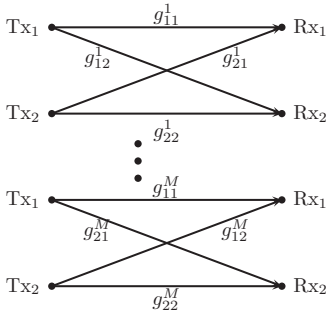

Fig. 1. Case study considered in Sec. III: an interference channel with 2 transmitters (Txs), 2 receivers (Rxs), and $M \geq 1$ non-overlapping frequency bands. One feature of the retained model is that both power allocation policies and channel gains g_{ij}^m are assumed to lie in finite discrete sets.

II. Optimization problem analysis

Evaluating the limiting performance in terms of average payoff amounts to characterizing the set of possible values for

978-3-901882-63-0/14/$31.00 ©2014 IEEE

$\mathbb{E}[u]$ under the information structure assumed in this paper. The expected payoff can be written as:

$$\mathbb{E}[u] = \sum_{(x_0,x_1,x_2)\in\mathcal{X}_0\times\mathcal{X}_1\times\mathcal{X}_2} q(x_0,x_1,x_2)u(x_0,x_1,x_2) \quad (1)$$

$$= \sum_{(x_0,x_1,x_2)\in\mathcal{X}_0\times\mathcal{X}_1\times\mathcal{X}_2} q_{X_0}(x_0)q(x_1,x_2|x_0)u(x_0,x_1,x_2) \quad (2)$$

where $q \in \Delta(\mathcal{X}_0\times\mathcal{X}_1\times\mathcal{X}_2)$, $\Delta(\cdot)$ standing for the unit simplex over the set under consideration, q_{X_0} is the marginal law of the random state and is considered fixed. The fact that the two agents can only coordinate through the considered information structure imposes a constraint on the average performance which translates into a constraint on q for the expected payoff [1][2]:

$$I_q(X_0;X_2) - H_q(X_1|X_0,X_2) \leq 0 \quad (3)$$

where, for any two random variables $(X,Y) \in (\mathcal{X}\times\mathcal{Y})$ with joint law $q(\cdot,\cdot)$:

- $H_q(X|Y)$ is the conditional entropy of X given Y defined by:

$$H_q(X|Y) = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} q(x,y)\log_2\frac{q(x,y)}{q_Y(y)} \quad (4)$$

  where $q_Y(\cdot)$ is obtained by marginalization of the joint distribution $q(\cdot,\cdot)$;
  One can note that the entropy of X is simply:

$$H_q(X) = -\sum_{x\in\mathcal{X}} q_X(x)\log_2 q_X(x) \quad (5)$$

- $I_q(X;Y)$ denotes the mutual information between X and Y, defined by:

$$I_q(X;Y) = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} q(x,y)\log_2\frac{q(x,y)}{q_X(x)q_Y(y)} \quad (6)$$

Reference [2] provides a clear interpretation of this constraint. Essentially, the first term can be seen as a rate-distortion term while the second term can be seen as a limitation in terms of communication medium capacity. To state the optimization problem which characterizes the limiting performance in terms of expected payoff, a few notations are in order. We denote the cardinality of the set $\mathcal{X}_i$, $i \in \{0,1,2\}$ as: $|\mathcal{X}_i| = n_i < \infty$. For the sake of simplicity and without loss of generality, we consider $\mathcal{X}_i$ as a set of indices $\mathcal{X}_i = \{1,...,n_i\}$. Additionally, we introduce the vector of payoffs (associated with the function u defined earlier) $w = (w_1,w_2,...,w_n) \in \mathbb{R}^n$ with $n = n_0n_1n_2$ and assume, without loss of generality, that $\Pr[X_0 = j] = \alpha_j > 0$ for all $j \in \mathcal{X}_0 = \{1,\dots,n_0\}$, with $\sum_{j=1}^{n_0}\alpha_j = 1$. The indexation of w and therefore the vector $q = (q_1,q_2,...,q_n)$ is chosen according to a lexicographic order. This is illustrated through Tab. I. This choice simplifies the analysis of the optimization problem which is stated next.

The Information Constraint (3) can be re-written as:

$$ic(q) \triangleq I_q(X_0;X_2) - H_q(X_1|X_0,X_2) \quad (7)$$

$$= H_q(X_0) + H_q(X_2) - H_q(X_0,X_1,X_2) \quad (8)$$

Index (i)	X_0	X_1	X_2
1	1	1	1
2	1	1	2
⋮	⋮	⋮	⋮
n_2	1	1	n_2
n_2+1	1	2	1
⋮	⋮	⋮	⋮
$2n_2$	1	2	n_2
⋮	⋮	⋮	⋮
$n_2(n_1-1)+1$	1	n_1	1
⋮	⋮	⋮	⋮
n_1n_2	1	n_1	n_2
⋮	⋮	⋮	⋮
$n_1n_2(n_0-1)+1$	n_0	1	1
⋮	⋮	⋮	⋮
$n_0n_1n_2$	n_0	n_1	n_2

TABLE I. CHOSEN INDEXATION FOR THE PAYOFF VECTOR w AND DISTRIBUTION VECTOR q. BOLD LINES DELINEATE BLOCKS OF SIZE n_1n_2 AND EACH BLOCK CORRESPONDS TO A GIVEN VALUE OF THE RANDOM STATE X_0.

With our notation, we have:

$$H_q(X_0) = -\sum_{i=1}^{n_0}\left[\Big(\sum_{j=1+(i-1)n_1n_2}^{in_1n_2} q_j\Big)\log_2\Big(\sum_{j=1+(i-1)n_1n_2}^{in_1n_2} q_j\Big)\right] \quad (9)$$

$$H_q(X_2) = -\sum_{i=1}^{n_2}\left[\Big(\sum_{j=0}^{n_0n_1-1} q_{i+jn_2}\Big)\log_2\Big(\sum_{j=0}^{n_0n_1-1} q_{i+jn_2}\Big)\right] \quad (10)$$

and

$$-H_q(X_0,X_1,X_2) = \sum_{i=1}^{n_0n_1n_2} q_i\log_2 q_i \quad (11)$$

Thus, the optimization problem of interest consists of finding the best joint distribution(s) q (i.e., the best correlation between the agent's actions and the random state) and is as follows:

$$\begin{aligned}
\min\quad & -\mathbb{E}_q[w] = -\sum_{i=1}^{n_0n_1n_2} q_iw_i \\
\text{s.t.}\quad & -1 + \sum_{i=1}^{n_0n_1n_2} q_i = 0 \\
& -\alpha_i + \sum_{j=1+(i-1)n_1n_2}^{in_1n_2} q_j = 0, \qquad \forall i \in \{1,\dots,n_0\} \\
& -q_i \leq 0, \qquad \forall i \in \{1,2,\dots,n_0n_1n_2\} \\
& ic(q) \leq 0
\end{aligned} \quad (12)$$

The first and third constraints imposes that q has to be a probability distribution. The second constraint imposes that the marginal of q with respect to x_1 and x_2 has to coincide with

the distribution of the random state which is fixed. The fourth constraint is the information-theoretic constraint (3).

To solve the optimization problem (12) we will apply the Karush Kuhn Tucker (KKT) necessary conditions for optimality [10]. For this purpose, we first verify that strong duality holds. This can be done e.g., by proving that Slater's constraint qualification conditions are met. Namely, there exists a strictly feasible point for (12) and that (12) is a convex problem. First, by specializing Lemma 1 in [2] in the case of perfect observation, we know that (3) defines a convex set. Since the cost function and the other constraints of the problem are affine, the problem is then convex; as a consequence. KKT conditions are also sufficient for optimality. The existence of a feasible point is stated in the next proposition.

Proposition 1. *There exists a strictly feasible distribution* $q^+ \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ *for the optimization problem* (12)*.*

Proof: First, choose a triplet of random variables (X_0, X_1, X_2) which are independent. That is, we consider a joint distribution q^+ which is of the form $q^+(x_0, x_1, x_2) = q^+_{X_0}(x_0) q^+_{X_1}(x_1) q^+_{X_2}(x_2)$. Second, one can always impose a full support condition to the marginals $q^+_{X_1}$ and $q^+_{X_2}$ (i.e., $\forall x_i, q^+_{X_i}(x_i) > 0$); $q^+_{X_0} \equiv q_{X_0}$ has a full support by assumption. Therefore, for the distribution $q^+(x_0, x_1, x_2)$ to be strictly feasible, it remains to be checked that the information-theoretic constraint is active. And this is indeed the case since:

$$\begin{aligned} I_q(X_0; X_2) - H_q(X_1|X_0, X_2) &= 0 - H_q(X_1|X_0, X_2) && (13)\\ &= -H_q(X_1) && (14)\\ &< 0 && (15)\end{aligned}$$

where: (13) and (14) come from the independence hypothesis between X_0, X_1, and X_2; (15) comes from the positiveness of the entropy and the fact that every $q^+(x_0, x_1, x_2)$ (and thus every $q^+_{X_1}(x_1)$) is strictly positive. ∎

Following the previous considerations, KKT conditions can be applied. The Lagrangian function can be written as:

$$\begin{aligned}\mathcal{L}(q, \mu, \mu_0, \lambda, \lambda_{\mathrm{IC}}) = &- \sum_{i=1}^{n_0 n_1 n_2} (w_i q_i + \lambda_i q_i)\\ &+ \mu_0 \left[\sum_{i=1}^{n_0 n_1 n_2} q_i - 1 \right] + \sum_{i=1}^{n_0} \mu_i \left[\sum_{j=1+(i-1)n_1 n_2}^{i n_1 n_2} q_j - \alpha_i \right]\\ &+ \lambda_{\mathrm{IC}} \cdot ic(q) \qquad (16)\end{aligned}$$

where $\lambda = (\lambda_1, ..., \lambda_{n_0 n_1 n_2})$, $\mu = (\mu_1, ..., \mu_{n_0})$, and IC stands for information-theoretic constraint. We have the following partial derivatives for the information constraint:

$$\begin{aligned}\frac{\partial ic}{\partial q_i}(q) = \Bigg[&- \sum_{k=1}^{n_0} \mathbb{1}_{\{1+(k-1)n_1 n_2 \le i \le (k) n_1 n_2\}}\\ &\qquad * \log_2 \Big(\sum_{j=1+(k-1)n_1 n_2}^{k n_1 n_2} q_j \Big)\\ &- \sum_{k=1}^{n_2} \mathbb{1}_{\{i \in \{k, k+n_2, \ldots, k+(n_0 n_1 - 1)n_2\}\}} \log_2 \Big(\sum_{j=0}^{n_0 n_1 - 1} q_{k+jn_2} \Big)\\ &+ \log_2 q_i - 1 \Bigg] \qquad \forall i \in \{1, 2, \ldots, n_0 n_1 n_2\} \qquad (17)\end{aligned}$$

Other terms of the Lagrangian are easy to derive. KKT conditions follow:

$$\begin{aligned}&\frac{\partial \mathcal{L}}{\partial q_i} = -w_i - \lambda_i + \mu_0 + \sum_{j=1}^{n_0} \mu_j \mathbb{1}_{\{1+n_1 n_2 (j-1) \le i \le j n_1 n_2\}}\\ &+ \lambda_{\mathrm{IC}} \left[\frac{\partial ic}{\partial q_i}(q) \right] = 0 \qquad \forall\, i \in \{1, 2, \ldots, n_0 n_1 n_2\} \qquad (18)\\ &\lambda_i \ge 0 \qquad \forall\, i \in \{1, 2, \ldots, n_0 n_1 n_2\} \qquad (19)\\ &\lambda_{\mathrm{IC}} \ge 0 \qquad (20)\\ &\lambda_i q_i = 0 \qquad \forall\, i \in \{1, 2, \ldots, n_0 n_1 n_2\} \qquad (21)\\ &\lambda_{\mathrm{IC}} ic(q) = 0 \qquad (22)\end{aligned}$$

where $\mathbb{1}_{\{\cdot\}}$ is the indicator function and $i(q)$ is the inequality constraint function associated with the information-theoretic constraint (3). By inspecting the KKT conditions, the following proposition can be proved.

Proposition 2. *If there exists a permutation such that the payoff vector* w *can be strictly ordered, then any optimal solution of (12) is such that the information-theoretic constraint is active i.e.,* $\lambda_{\mathrm{IC}} > 0$*.*

Proof: We proceed by contradiction. Assume that the payoff vector can be strictly ordered and that the constraint is not active for solutions under consideration, that is, $\lambda_{\mathrm{IC}} = 0$.

First, consider possible solution candidates q which have two or more non-zero components per block of size $n_1 n_2$ which is associated with a given realization x_0 of the random state (see Tab. I) . Since there exists a pair of distinct indices (j, k) such that $q_j > 0$, $q_k > 0$, we have that $\lambda_j = 0$, $\lambda_k = 0$. This implies that, through the gradient conditions of the KKT conditions, $w_j = w_k$ which contradicts the fact that payoffs are strictly ordered.

Second, consider possible solution candidates q which have only one non-zero component per block associated with x_0 (see Tab. I). This implies that $H_q(X_0, X_1, X_2) = H_q(X_0) = H(X_0)$, which means that $H_q(X_0) + H_q(X_2) > H_q(X_0, X_1, X_2)$, whenever $H_q(X_2) > 0$. This means that the constraint is violated and therefore the considered candidates are not feasible. Now, if $H_q(X_2) = 0$, we see that the Information constraint is active which contradicts again the starting assumption . ∎

Proposition 2 is especially useful for wireless communications when the state is given by the overall channel. Due

to channel randomness, the most common scenario is that the payoffs associated with the channel realizations are distinct. For this reason, we will assume such a setting in this paper and thus that $\lambda_{\mathrm{IC}} > 0$. If $\lambda_{IC} > 0$, we have the following:

- We can not have $\lambda_i > 0$ for one or more $i \in \{1, 2, \ldots, n_0 n_1 n_2\}$. Indeed, if for example $\lambda_i > 0$, then $q_i = 0$, which implies $\log_2(q_i) = -\infty$ and (18) can not be satisfied.
- However, if one of the q_i's equals 0, and $q_k = 0$ for all k such that $k[n_2] = i[n_2]$ (where $[x]$ stands for modulo x), then the λ_{IC} component equals $\lim_{x\to 0} \frac{x}{n_0 n_1 x}$ and does not go to $-\infty$. This case cannot be discarded, but it can be said that X_2 is deterministic in such a case.

Summarizing our analysis, the only possible cases are:

- $\lambda_{\mathrm{IC}} > 0$, and exactly one λ_i for each block (corresponding to a particular state of nature) are non-zeros, and they have to be associated with the same action of X_2 (X_2 has to be deterministic). In this case there is no communication, and the optimal strategies are trivial. Therefore we shall not be discussing this case henceforth.
- The only relevant case is:

$$\begin{aligned} \lambda_i &= 0 \qquad \forall\, i \in \{1, 2, \ldots, n_0 n_1 n_2\} \\ \lambda_{\mathrm{IC}} &> 0 \end{aligned}$$

For the latter case, KKT conditions become:

$$\frac{\partial \mathcal{L}}{\partial q_i} = -w_i + \mu_0 + \left(\sum_{j=1}^{n_0} \mu_j \mathbb{1}_{\{1+n_1 n_2 (j-1) \le i \le j n_1 n_2\}} \right) + \lambda_{\mathrm{IC}} \left[\frac{\partial ic}{\partial q_i}(q) \right] = 0 \qquad \forall\, i \in \{1, 2, \ldots, n_0 n_1 n_2\} \quad (23)$$

$$\lambda_i = 0 \qquad \forall\, i \in \{1, 2, \ldots, n_0 n_1 n_2\} \quad (24)$$

$$\lambda_{\mathrm{IC}} > 0 \quad (25)$$

$$i(q) = 0. \quad (26)$$

Now that we have proved some useful results about the structure of optimal solutions of (12), a natural question is whether the optimal solution is unique, which is the purpose of the next proposition.

Proposition 3. *If there exists a permutation such that the payoff vector w can be strictly ordered, the optimization problem (12) has a unique solution.*

Proof: We know, by Prop. 2, that $\lambda_{\mathrm{IC}} > 0$ for any optimal solution. It turns out that, if $\lambda_{\mathrm{IC}} > 0$, the Lagrangian of (12) is a strictly convex function w.r.t. the vector q. Indeed, the optimization spaces are compact and convex, and the Lagrangian is the sum of linear functions and a strictly convex function $i(q)$.
It remains to show that $ic : q \mapsto I_q(X_0; X_2) - H_q(X_1|X_0, X_2)$ is strictly convex over the set of distributions $q \in \Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ that verify $q_{X_0} := \sum_{(x_1,x_2)} q(x_0, x_1, x_2) = \rho(x_0)$ with ρ fixed.

The first term $I_q(X_0; X_2)$ is a convex function of $q_{X_2|X_0}$ for fixed q_{X_0}. For the second term, let $\lambda_1 \in [0, 1]$, $\lambda_2 = 1 - \lambda_1$, $(q^1, q^2) \in (\Delta(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2))^2$ and $q = \lambda_1 q^1 + \lambda_2 q^2$. We have that:

$$H_q(X_1|X_0, X_2) = -\sum_{x_0,x_1,x_2} \left(\sum_{i=1}^{2} \lambda_i q^i(x_0, x_1, x_2) \right) \cdot \log \left[\frac{\sum_{i=1}^{2} \lambda_i q^i(x_0, x_1, x_2)}{\sum_{i=1}^{2} \lambda_i q^i_{X_2}(x_2)} \right] \quad (27)$$

$$> -\sum_{x_0,x_1,x_2} \sum_{i=1}^{2} \lambda_i q^i(x_0, x_1, x_2) \log \left[\frac{\lambda_i q^i(x_0, x_1, x_2)}{\lambda_i q^i_{X_2}(x_2)} \right] \quad (28)$$

$$= -\sum_{i=1}^{2} \lambda_i \sum_{x_0,x_1,x_2} q^i(x_0, x_1, x_2) \log \left[\frac{q^i(x_0, x_1, x_2)}{q^i_{X_2}(x_2)} \right] \quad (29)$$

$$= \lambda_1 H_{q^1}(X_1|X_0, X_2) + \lambda_2 H_{q^2}(X_1|X_0, X_2) \quad (30)$$

where (28) comes from the log sum inequality [9], with:

$$a_i = \lambda_i q^i(x_0, x_1, x_2) \quad (31)$$

and

$$b_i = \lambda_i q^i_{X_2}(x_2) \quad (32)$$

for $i = 1, 2$ and for all x_0, x_1, x_2 such that $q^i_{X_2}(x_2) > 0$.

The inequality is strict because $\frac{a_1}{b_1} \neq \frac{a_2}{b_2}$, since we have assumed that q^1 and q^2 distinct. ■

The uniqueness property for the optimization problem is particularly useful in practice since it means that any converging numerical procedure to find an optimal solution will lead to the unique global minimum.

III. Distributed power allocation case study

A. Case study description

We now consider the specific problem of power allocation over $M-$band interference channels with two transmitter-receiver pairs. Transmissions are time-slotted and, on each time-slot, transmitter $i \in \{1, 2\}$ has to choose a power allocation vector in the following set of actions:

$$\mathcal{P}_i = \left\{ \frac{P_{\max}}{\ell} e_\ell : \ell \in \{1, \ldots, M\}, \; e_\ell \in \{0, 1\}^M, \sum_{i=1}^{M} e_\ell(i) = \ell \right\} \quad (33)$$

where $P_{\max}$ is the the power budget available at a transmitter. Each channel is assumed to lie in a discrete set $\Gamma = \{g_1, ..., g_S\}$, $S \geq 1$, $g_s \geq 0$ for $s \in \{1, ..., S\}$. Therefore, if one denotes by g^m the vector of four channel gains corresponding to the band $m \in \{1, ..., M\}$, then $g^m \in \Gamma^4$ and the global channel state $g = [g^1, ..., g^M]$ lies in $\mathcal{G} = \Gamma^{4M}$ whose cardinality is S^{4M}. As it is always possible to find a one-to-one mapping between $\mathcal{P}_i$, $i \in \{1, 2\}$, (resp. $\mathcal{G}$) and $\mathcal{X}_i$ (resp. $\mathcal{X}_0$) as defined in Sec. II, the results derived therein can be applied here. Lastly, for a given time-slot, the instantaneous or

the stage payoff function which is common to the transmitters is chosen to be:

$$u : \left| \begin{array}{l} \mathcal{G} \times \mathcal{P}_1 \times \mathcal{P}_2 \to \mathbb{R}^+ \\ (g, p_1, p_2) \mapsto \sum_{i=1}^{2} \sum_{m=1}^{M} B_m \log_2 \left(1 + \frac{g_{ii}^m p_i^m}{\sigma^2 + g_{-ii}^m p_{-i}^m}\right) \end{array} \right. \tag{34}$$

where p_i is the power allocation chosen by transmitter i on the current time-slot whose channel state is g, σ^2 is the noise variance, B_m is the bandwidth of band m, p_i^m the power transmitter i allocates to band i,$-i$ stands for the transmitter other than i.

B. Simulation setup

In this section, specific values for the parameters which are defined in the preceding section are chosen, in particular to make the interpretations relatively easy. We assume $M = 2$ bands and therefore that the transmitters have three actions: $\mathcal{P}_i = P_{\max}\left\{(0,1),(1,0),(\frac{1}{2},\frac{1}{2})\right\}$ for $i \in \{1,2\}$. As [11] we assume the first band to be protected ($g_{12}^1 = g_{21}^1 = 0$) whereas the second band corresponds to a general single-band interference channel. The other channel gains are chosen as follows:

$$g_{ii}^1 \in \{0.1, 1.9\}, \quad i \in \{1,2\} \tag{35}$$

$$g_{ij}^2 \in \{0.15, 1.85\}, \quad (i,j) \in \{1,2\}. \tag{36}$$

We suppose that each g_{ij}^k, $k = 1,2$ is i.i.d. and Bernouilli distributed $g_{ij}^k \sim \mathcal{B}(\pi_{ij}^k)$ with $P(g_{ii}^1 = 0.1) = \pi_{ii}^1$ and $P(g_{ij}^2 = 0.15) = \pi_{ij}^2$. We define SNR[dB]$= 10\log_{10}\left(\frac{P_{\max}}{\sigma^2}\right)$, and we consider two regimes for the second band: a high interference regime (HIR), defined by $(\pi_{11}^2, \pi_{12}^2, \pi_{21}^2, \pi_{22}^2) = (0.5, 0.1, 0.1, 0.5)$ and a low interference regime (LIR) defined by $(\pi_{11}^2, \pi_{12}^2, \pi_{21}^2, \pi_{22}^2) = (0.5, 0.9, 0.9, 0.5)$. For the first band, we take $\pi_{11}^1 = \pi_{22}^1 = 0.2$. One can see that our choice of parameters indeed define a high interference regime: $P((g_{ij}^2 | i \neq j) = 1.85) = 1 - 0.1 = 0.9$, thus creating high interference due a high probability for a greater value of $(g_{ij}^2 | i \neq j)$. The similar intuition holds for low interference regime. Three power allocation policies will be considered:

- The costless communication case, where both transmitters knows the state beforehand and can reach the maximum payoff at every stage;
- The (information-constrained) optimal policy (OP) corresponding to the optimal solution of the optimization problem (12);
- The blind policy (BP), where transmitters don't know anything about channel gains and always choose to put half of their power in each band: $p_1 = p_2 = P_{\max}(\frac{1}{2}, \frac{1}{2})$ at every stage.

Fig. 2 represents the gain allowed by asymmetric coordination w.r.t. the case where the transmitters always use the uniform power allocation policy (BP). This gain can be as high as 40% for the considered range of SNR. It is seen that the gains are particularly significant when the interference is high (the two top red curves) and in the low and high SNR regimes (red and blue curves on the left and right sides). The first observation translates the intuition that the higher the interference level the stronger is the gain brought by coordination. The second can be understood as follows: In the high SNR regime, the transmission rate over the non-protected band is interference limited and bounded and it is better to allocate the power to the protected band which allows an arbitrarily large rate as the SNR grows large. This explains why allocating uniformly the power becomes more and more suboptimal as the SNR increases. In the low SNR regime, essentially the interference becomes negligible and the best power allocation policies roughly correspond to water-filling over the available channels. At low SNR, the best water-filling policy is to use the best band and not to allocate power uniformly, which explains the gap between the coordinated policies and uniform power allocation. Our explanations are sustained by Fig. 3, which shows the probability that a transmitter uses a given power allocation vector. For instance, at low SNR, the dominant actions for both transmitters is to use the protected band. It can be noticed that transmitter 1 has also to convey information to transmitter 2 (i.e., ensuring that the entropy of X_1 is not too small), which is why he cannot use the protected band as often as transmitter 2. One also notices in Fig. 3 that the probability of the action $(0,1)$ (using the shared band) is zero from lower SNR values for transmitter 2 than for transmitter 1. This can be explained by the fact that the higher the power available for both transmitters, the higher the interference in the non-protected band. However, transmitter 1 still chooses to play this action as it has knowledge of channel gains and can use the interference band to improve the common utility. The same argument stands for Fig. 4. At last, Fig. 4 shows the influence of the bandwidths on the power allocation policies. Not surprisingly, the higher the bandwidth of the protected band, the more often it is used, and conversely for the non-protected band. Concerning the uniform policy, it is seen that transmitter uses it more frequently, although channel conditions are similar, which translates again the need for transmitter 1 to convey information.

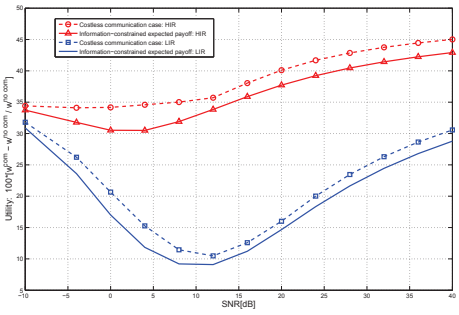

Fig. 2. Relative gain in terms of expected payoff ("OP/BP - 1" in [%]) vs SNR[dB] obtained with the Optimal policy (OP) (with and without communication cost) when the reference policy is to put half of the power on each band (BP). Red curves correspond to the HIR, and blue curve to the LIR. $B_1 = B_2 = 10$MHz.

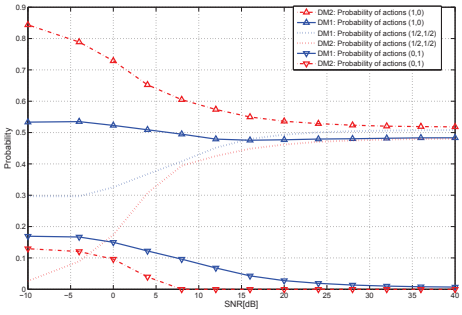

Fig. 3. Marginal probability distributions $q_{X_1}(\cdot)$ $q_{X_2}(\cdot)$ of transmitter 1 and transmitter 2 for the optimal policy vs SNR[dB] for the High Interference Regime. $B_1 = B_2 = 10$MHz.

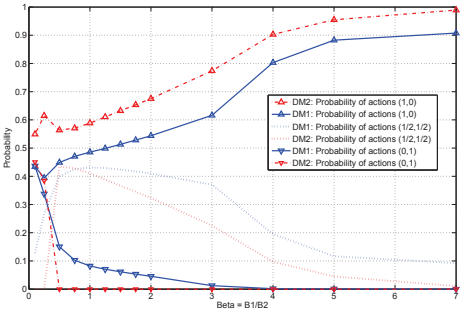

Fig. 4. Marginal probability distributions $q_{X_1}(\cdot)$ $q_{X_2}(\cdot)$ of transmitter 1 and transmitter 2 for the optimal policy vs β for the optimal policy for the High Interference Regime, where $\beta = \frac{B_1}{B_2}$. For this simulation, we chose SNR=10[dB].

IV. Conclusion

This work clearly illustrates the potential benefit of the proposed approach, by embedding coordination information into the power allocation levels, relative gains as high as 40% can be obtained w.r.t. the uniform power allocation policies. In this work, the embedded information is a distorted version of the channel state but the proposed approach is much more general: information about the state of queue, a battery, etc, could be considered; other types of policies might be considered to encode information e.g., channel selection policies, transmit power levels. The study of generalized versions of this problem, such as the case of imperfect monitoring, or continuous power allocation, will be provided in future works. This work however indicates the high potential of such an approach for team optimization problems. More importantly, it gives an optimization framework to analyze performance limits for problems with implicit communication.

Acknowledgment

Auhtors of this paper and particularly A. Agrawal would like to thank project LIMICOS - ANR-12-BS03-0005 for financing the project.

References

[1] O. Gossner, P. Hernandez, and A. Neyman, "Optimal use of communication resources," *Econometrica*, vol. 74, no. 6, pp. 1603–1636, Nov. 2006.

[2] B. Larrousse and S. E. Lasaulce, "Coded power control: Performance analysis," in *2013 IEEE International Symposium on Information Theory (ISIT'2013)*, Jul. 2013.

[3] I. Khayutin, "Communication with unobservable constraints," in *Research report*, Dec. 2007.

[4] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed multiuser power control for digital subscriber lines," *IEEE Journal on selected areas in Communications*, vol. 20, no. 5, pp. 1105 – 1115, June 2002.

[5] Y. Xing and R. Chandramouli, "Stochastic learning solution for distributed discrete power control game in wireless data networks," *IEEE/ACM Trans. Networking*, vol. 16, no. 4, pp. 932–944, Aug. 2008.

[6] E. V. Belmega, H. Tembine, and S. Lasaulce, "Learning to precode in outage minimization games over MIMO interference channels," in *IEEE Proc. of the Asilomar Conference on Signals, Systems, and Computer*, Nov. 2010.

[7] L. Rose, S. Lasaulce, S. M. Perlaza, and M. Debbah, "Learning equilibria with partial information in wireless networks," *IEEE Communication Magazine*, vol. 49, no. 8, pp. 136–142, Aug. 2011, special Issue on Game Theory for Wireless Networks.

[8] S. Lasaulce and H. Tembine, *Game Theory and Learning for Wireless Networks : Fundamentals and Applications*, Elsevier, Ed. Academic Press, Aug. 2011.

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 2006.

[10] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[11] R. Mochaourab and E. Jorswieck, "Resource allocation in protected and shared bands: uniqueness and efficiency of nash equilibria," in *Fourth International ICST Conference on Performance Evaluation Methodologies and Tools*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Oct. 2009, p. 68.

E.2.3 IEEE-ICASSP 2014

CRAWFORD-SOBEL MEET LLOYD-MAX ON THE GRID

B. Larrousse, O. Beaude, and S. Lasaulce

L2S (CNRS-Supélec-Univ. Paris Sud 11), 3 Rue Joliot-Curie, 91192 Gif-sur-Yvette, France
{ larrousse, beaude, lasaulce } @lss.supelec.fr

ABSTRACT

The main contribution of this work is twofold. First, we apply, for the first time, a framework borrowed from economics to a problem in the smart grid namely, the design of signaling schemes between a consumer and an electricity aggregator when these have non-aligned objectives. The consumer's objective is to meet its need in terms of power and send a request (a message) to the aggregator which does not correspond, in general, to its actual need. The aggregator, which receives this request, not only wants to satisfy it but also wants to manage the cost induced by the residential electricity distribution network. Second, we establish connections between the exploited framework and the quantization problem. Although the model assumed for the payoff functions for the consumer and aggregator is quite simple, it allows one to extract insights of practical interest from the analysis conducted. This allows us to establish a direct connection with quantization, and more importantly, to open a much more general challenge for source and channel coding.

Index Terms— Best-response, Nash equilibrium, signal quantization, smart grid, strategic information transmission.

1. INTRODUCTION

In today's electricity networks, the energy production is mainly driven by the consumer's demand. However, in the smart grid it will be more common that the consumer will have to adapt its consumption to production e.g., when an erasure mechanism is implemented or when the energy source is a solar/wind farm. Obviously, the consumer and aggregator (i.e., the entity which takes the decision to which extent to meet the demand) will have diverging objectives in general. As a consequence, it might happen that the consumer reports a demand which is higher than the actual need to be effectively satisfied. As, in practice, the request of the consumer (a factory, a house, an EV's owner, etc.) in terms of needed power and the decision to which extent to deliver it by the aggregator (a utility company, a distribution network's operator, etc) will quite often result from (automated) procedures implemented by machines, an important engineering problem appears: How to design a point-to-point communication system where the transmitter (or coder) and receiver (or decoder) have diverging objectives? Indeed, the classical paradigm in communication systems [1], assumes that the coder and decoder have a common objective (e.g., to minimize the distortion or block error rate). When the coder and decoder have non-aligned objectives, the problem of (source/channel) coding needs to be revisited. In the present paper, we will only make a small step into the direction of answering the aforementioned fundamental question. Nonetheless, the work reported here has the merit to bridge an obvious gap between the economics literature and the one of signal processing and communications.

Specifically, we consider an aggregator whose objective is to satisfy the consumer but also to minimize the operating cost induced by the distribution network. More precisely, the cost of the distribution network is chosen to be the residential transformer ageing[1]. On the other hand, the consumer's ultimate objective is to obtain an amount of power (or energy) as close as possible to its actual need. Based on a signal/message received from the consumer about its need in terms of power, the aggregator eventually decides the amount of power effectively allocated to the consumer. One of the purposes of this paper is to construct a signaling scheme from the consumer to the aggregator which would allow them to reach a consensus or equilibrium about how to communicate in practice (based on a suited communication standard). It turns out that, by considering a simple but realistic model for the aggregator and consumer costs, the problem to be solved is a game whose formulation is related to the problem of strategic information transmission in economics [5] and the one of quantization. Indeed, the problem of strategic information transmission has been introduced in [5] and developed in economics[2] (see e.g., [6] for a recent survey) but not penetrated engineering yet up to a few exceptions [7][8], which do not consider neither the smart grid application nor the connections with coding/quantization.

The paper is organized as follows. The problem is formulated in Sec. 2. The signaling scheme is determined in Sec. 3

[1]Note that, in France, for instance, there are about 700 000 residential transformers, which shows the importance of managing transformers ageing; see e.g., [2][3][4] for more motivations.

[2]A typical example in economics is the interaction between a recruiting officer and a job seeker (the latter has to reveal more or less information about his state, i.e. his capabilities or expectations in terms of salary, while the former has to decide about the salary level).

as well as its main properties of practical interest. Sec. 4 provides numerical results to illustrate the derivations of Sec. 3. Sec. 5 concludes the paper.

2. PROBLEM FORMULATION

Fig. 1 provides several key aspects of the considered problem. We consider a consumer whose objective is to obtain an allocated power which is as close as possible to a desired level denoted by $s \geq 0$. For this purpose, the consumer sends a message $m \in \{1, 2, ..., M\}$ ($M < +\infty$) to the aggregator through a perfect communication channel. Based on the received message, the aggregator effectively provides an amount of power which is denoted by $a \geq 0$. Without loss of generality, it is assumed that $(a, s) \in [0, 1]^2$. One way of mathematically formulating the objective of the consumer, is to consider that he aims at maximizing the following payoff function

$$u_{\text{C}}(s, a) = -(s - a)^2 + K \tag{1}$$

where $K \in \mathbb{R}$ is a constant. With such a model, the consumer both aims at meeting its need in terms of power but also at not exceeding the desired power level, which might for instance induce some unnecessary monetary expenses. This model can also be very well justified when s is interpreted as a desired quantity of energy e.g., for recharging a battery (see, e.g., [9]). Note that here, for the sake of simplicity, we implicitly assume that the energy need corresponds to a need in terms of load or power, which is very realistic when the consumer obtains a constant power transfer rate; relaxing this assumption can be considered as a possible extension of this work. On the other hand, the aggregator's payoff function is assumed to be the weighted sum of the consumer's payoff and a payoff function related to an operating cost induced by the grid:

$$\begin{aligned} u_{\text{A}}(s, a) &= u_{\text{C}}(s, a) + u_{\text{grid}}(s, a) \\ &= -(s - a)^2 + K - be^a \end{aligned} \tag{2}$$

where $b \geq 0$ represents a weight which translates the importance of the component associated with the grid. More precisely, the grid component represents a good model of the ageing acceleration factor of a (residential) transformer (see e.g., [10] which justifies why the ageing is accelerated exponentially when operating above its nominal load). In the context of strategic information transmission in economics [5], the parameter b is interpreted as a bias which quantifies the divergence of interests between the decision-makers which are the consumer and aggregator here.

One of the contributions of this paper is precisely to inspire from the original framework of [5] to design a good/consensus/equilibrium signaling scheme between the consumer and aggregator namely, to determine a good signaling scheme in presence of diverging interests between the coder and decoder. First, the consumer should map its knowledge about its actual power need s into the message sent to the aggregator m, which amounts to determining a coding function f defined by:

$$f : \left| \begin{array}{ccc} [0,1] & \to & \{1, 2, ..., M\} \\ s & \mapsto & m \end{array} \right. . \tag{3}$$

Second, the aggregator has to perform the decoding operation by implementing:

$$g : \left| \begin{array}{ccc} \{1, 2, ..., M\} & \to & [0,1] \\ m & \mapsto & a \end{array} \right. . \tag{4}$$

As a first comment note that f and g are deterministic mappings instead of conditional probabilities $q(m|s)$ and $r(a|m)$; this choice does not induce any loss in terms of expected payoff because u_{A} and u_{C} are concave. If $b = 0$ and the power need s is seen as the realization of a random variable whose distribution $p(s)$ is effectively known to the coder and decoder (this corresponds to a particular scenario in terms of beliefs), the problem of determining f and g can be seen as an instance of a scalar quantization problem which is itself a special case of lossy source coding [11]. But, in general $b > 0$ and, even if the distribution $p(s)$ is known to both the coder and decoder, the consequence of this simple difference is that the coder, knowing that the decoder has a different objective, will not maximize its expected payoff by revealing its actual need in terms of power. Rather, it will reveal only a degraded version of it and, this, even if M is infinite. As explained in the following section, in general, equilibrium signaling schemes only exploit a fraction of the number of available messages (or bits).

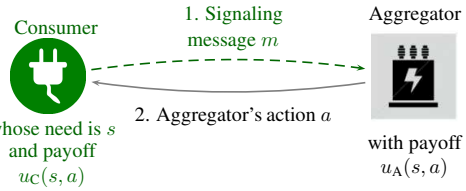

Fig. 1. The consumer (coder) has an actual need in terms of power s which is unknown to the aggregator. The consumer sends a message m to the aggregator (decoder). The aggregator then chooses an action a, which is the amount of power effectively allocated to the consumer. The key point is that the aggregator and consumer have *non-aligned* payoff functions.

3. PROPOSED SIGNALING SCHEME

3.1. Methodology. Connection with quantization

In the presence of decision-makers having different payoff functions and which can only control some variables of the

latter, the very meaning of optimality is unclear and the problem needs to be defined before being solved (see e.g., [12]). In this context, one important solution concept is the Nash Equilibrium (NE), which is a vector of strategies from which no decision-maker or player has anything to gain by changing his own strategy unilaterally. Here, we are in the presence of two players namely, the aggregator and consumer. The strategy of the consumer consists in choosing f, which corresponds to choosing a partition of the space of possible power needs i.e., $[0,1]$. With each interval is associated a message $m \in \{1,...,M\}$ intended for the aggregator. The strategy of the aggregator consists in choosing g to generate the action a, which can be interpreted as choosing a representative of the interval associated with the received message m; these intervals are denoted by $I_m = [s_m, s_{m+1}]$. Here, the connection with the quantization problem can be established. Typically, the quantization problem consists in minimizing the distortion $\mathbb{E}\left[(s-\widehat{s})^2\right]$ ($\widehat{s} = a$ in our setting), with respect to f and g. If f and g are optimized separately, the problem can be interpreted as a game where one player chooses f and the other chooses g. Since the cost functions are common and the number of message M is fixed, this defines a potential game [13][14]. In this type of games, it is known that the iterative procedure consisting in optimizing the cost/payoff function w.r.t. f for a fixed g, then to optimize it w.r.t. g for the updated f, and so on, converges to an NE. This procedure is called the sequential best-response (BR) dynamics in game theory, the BR of a player being the set-valued function which provides the set of strategies which maximize the payoff of this player for a given strategy for the other. The Lloyd-Max algorithm precisely implements this procedure and converges to an NE. Indeed, the intersection points between the players' BRs are precisely the NE of the game. In the next section, we determine the BRs in the considered setting in which players have different payoff functions.

3.2. Aggregator's best-response

When the aggregator receives a message m, it knows that the actual consumer's power need s is in the interval I_m but not its exact value. Therefore, in general, given the knowledge of the message, the aggregator has a certain belief about the power need, which is denoted by $\pi_{\mathrm{A}}(s|m)$. The aggregator best-responds to the message by maximizing the expected payoff that is,

$$a^\star(m) = \int_0^1 u_{\mathrm{A}}(a,s)\pi_{\mathrm{A}}(s|m)\mathrm{d}s. \tag{5}$$

Here, we assume that this belief is chosen to be a uniform probability distribution over the interval I_m which corresponds to the case where the aggregator has no statistical information at all about the consumer's power need; other scenarios in terms of belief are left as extensions. The following proposition provides the expression of the aggregator BR i.e., the best representative of the interval I_m.

Proposition 1. *Given a partition scheme f (or $m(s)$), the aggregator's best-response $a^\star(m)$ to a message m is:*

$$a^\star(m) = \begin{cases} \overline{s}_m - W\left(b\frac{e^{\overline{s}_m}}{2}\right) & \text{if } \overline{s}_m > \frac{b}{2} \\ 0 & \text{if } \overline{s}_m \leq \frac{b}{2} \end{cases} \tag{6}$$

where $\overline{s}_m = \frac{s_m+s_{m+1}}{2}$ and W is the Lambert W function[3].

The proof of this result is not provided here. It is constructed from arguments exploited in [5] which uses a useful property (called the single crossing condition) of the payoff functions. This property is as follows:

$$\forall i \in \{\mathrm{A},\mathrm{C}\}, \begin{cases} \forall s, & \exists a, \frac{\partial u_i}{\partial s}(s,a) = 0 \\ \forall (a,s), & \frac{\partial^2 u_i}{\partial s^2}(a,s) < 0 \\ \forall (a,s), & \frac{\partial^2 u_i}{\partial s \partial a}(a,s) > 0 \end{cases} . \tag{7}$$

The above result shows that it is possible to express the aggregator's best action (for a given message) in a simple way. The first term of the optimal action corresponds to what is called the centroid in quantization. The presence of the second term is precisely due to the fact that the coder and decoder have diverging interests. In the extreme case where $b \to 0$, the optimal action for the aggregator therefore corresponds to the centroid whereas the optimal action is simply 0 when $b \to \infty$.

3.3. Consumer's best-response

The consumer's strategy is to choose a partition of the power need space $[0,1]$ into intervals $I_1, I_2, ..., I_M$ with $I_m = [s_m, s_{m+1}]$. In contrast with a classical quantization problem, the number of messages (or bits) to be used to form the partition is not fixed and can be optimized by the consumer in order to maximize its expected payoff function for a given action. This feature constitutes an important technical difference which can be incorporated in a non-trivial manner. One key result of [5] which we re-exploit here is that the optimal partition structure can be built from an optimality condition called the arbitrage condition. Intuitively, if the consumer power need s is exactly s_m, then the consumer should be indifferent between sending the messages $m-1$ (associated with the interval $[s_{m-1}, s_m]$) and m (associated with the interval $[s_m, s_{m+1}]$). By exploiting this optimality condition and the single crossing condition (7), the following proposition can be proved.

Proposition 2. *Let $M_b^\star$ be the number of optimal partitions. For a given $L \leq M_b^\star$, the optimal partition for the consumer can be defined recursively as:*

$$\begin{cases} s_0^\star = 0 \\ s_{m+1}^\star = \phi_b(2s_m^\star - \phi_b^{-1}(s_{m-1}^\star + s_m^\star)) - s_m^\star, \\ \qquad\qquad 1 \leq m \leq L-1 \\ s_L^\star = 1 \end{cases} \tag{8}$$

[3] Some basics on the Lamber W function can be found here: http://mathworld.wolfram.com/LambertW-Function.html

where $\phi_b(x) = 2x + be^x$ and ϕ_b^{-1} its inverse function.

Prop. 1 and 2 completely define the signaling scheme for the problem under investigation. Indeed, when the consumer chooses a partition of the power need space according to Prop. 2 and the aggregator chooses the representatives according to Prop. 1 , we obtain an NE. The number $M_b^\star$ corresponds to the number of NE. This number can be obtained by choosing a certain $s_1 = b + \epsilon$ and determining the partition $s_2(s_1), s_3(s_1), ...$ through $s^\star_{m+1} = \phi_b(2s^\star_m - \phi_b^{-1}(s^\star_{m-1} + s^\star_m)) - s^\star_m$, and keeping the greatest integer m such that $s^\star_{m+1} < 1$. It can be shown [5] that an NE based on partition exploiting $L \in \{1, 2, ..., M_b^\star\}$ messages is better ex ante for both players than another NE which exploits $L' < L$ messages. If the bias is greater than a threshold β, $M_b^\star = 1$, which means that the consumer's message set is a singleton and no information is revealed to the aggregator. This threshold is important in practice because it allows a designer to know under which conditions a given signaling-based power production cannot be implemented. As a last comment on Prop. 1, note that when $b \to 0$, the recursive equation boils down to $s^\star_{m+1} = 2s^\star_m - s^\star_{m-1}$ i.e., $s^\star_m = \frac{s^\star_{m-1}+s^\star_{m+1}}{2}$. Therefore, a partition similar to uniform quantization is obtained.

4. NUMERICAL RESULTS

We assume that the consumer's power need is distributed uniformly over $[0, 1]$, $K = 1$, and we study the influence of the parameter b. In practice, b can be determined by the weight the aggregator puts on the transformer's cost but also physical parameters such as the ambient and hot-spot temperatures, and the nominal load [10]. Fig. 2 quantifies the relation between the bias and the maximum number of messages used at equilibrium (which also corresponds to the number of NE). Here, for $b \geq \beta$, with $\beta \sim 0.25$, the consumer does not reveal anything about its need in terms of power to the aggregator. When $b \to 0$, the number of messages becomes high and, will be limited, in practice, e.g, by the communication channel capacity. Fig. 3 shows the expected payoff obtained by the aggregator ("A") and consumer ("C") as a function of b in different scenarios. The three bottom curves correspond to the aggregator's payoff when: 1) "A" has access to no message; 2) "A" receives the message from "C" (equilibrium payoff); 3) "A" is given the perfect knowledge of the power need. The loss induced by the uncertainty on the power need corresponds to the gap between 1) and 3), which may be typically much higher for other payoff functions. The three top curves correspond to the consumer's payoff when: 4) "C" $a = s$; 5) "A" knows the actual power need; 6) "C" sends a message to "A" (equilibrium payoff).

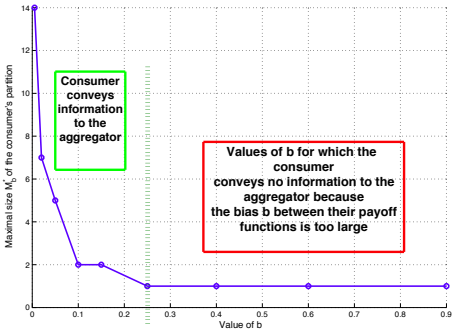

Fig. 2. Maximal number of messages $M_b^\star$ vs the bias b.

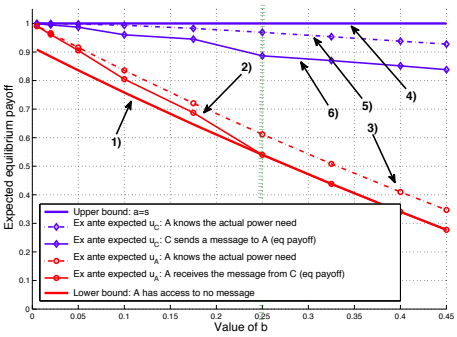

Fig. 3. Consumer's and aggregator's payoff vs the bias b.

5. CONCLUSION

Obviously, the model used in this work can be generalized in many respects e.g., in terms of payoff functions, in terms of beliefs about the source/need distribution, and also by considering sequence of actions instead of a single action. The framework which is exploited in this paper goes beyond the quantization aspects and the new connections established between [5] and quantization and opens new technical challenges which concern the general problem of source and channel coding when the coders and decoders have different performance criteria.

6. REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.

[2] V. Marano Q. Gong, S. Midlam-Mohler and G. Rizzoni, "Study of pev charging on residential distribution transformer life," *IEEE Transactions on Smart Grid*, 2011.

[3] K. D. McBee and M. G. Simoes, "Reducing distribution transformer losses through the use of smart grid monitoring," *North American Power Symposium (NAPS)*, pp. 1–6, 2009.

[4] Tran Q. T., X. Le Pivert, M. Saheli, and O. Beaude, "Stochastic approach to assess impacts of electric vehicles on the distribution network," in *Innovative Smart Grid Technologies (ISGT Europe), 2012 3rd IEEE PES International Conference and Exhibition on*, 2012, pp. 1–8.

[5] V. Crawford and J. Sobel, "Strategic information transmission," *Econometrica: Journal of the Econometric Society*, pp. 1431–1451, 1982.

[6] J. Sobel, "Signaling games," *Encyclopedia of Complexity and System Science, Springer, forthcoming*, 2007.

[7] V. Kavitha, E. Altman, R. El-Azouzi, and R. Sundaresan, "Opportunistic scheduling in cellular systems in the presence of noncooperative mobiles," *Information Theory, IEEE Transactions on*, vol. 58, no. 3, pp. 1757–1773, 2012.

[8] F. Mériaux, S. Valentin, S. Lasaulce, and M. Kieffer, "An energy-efficient power allocation game with selfish channel state reporting in cellular networks," in *Performance Evaluation Methodologies and Tools (VALUETOOLS), 2012 6th International Conference on*, 2012, pp. 290–297.

[9] V. Marano, S. Onori, Y. Guezennec, G. Rizzoni, and N. Madella, "Lithium-ion batteries life estimation for plug-in hybrid electric vehicles," in *Vehicle Power and Propulsion Conference, 2009. VPPC '09. IEEE*, 2009, pp. 536–543.

[10] IEC, *Loading guide for Oil-Immersed Power Transformers*, Number 60354. IEC, 2011.12.02.

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*, Wiley-Interscience, 2006.

[12] S. Lasaulce and H. Tembine, *Game Theory and Learning for Wireless Networks : Fundamentals and Applications*, Academic Press, 2011.

[13] D. Monderer and L. Shapley, "Potential games," *Games and Economic Behavior*, vol. 14, no. 1, pp. 124 – 143, 1996.

[14] M. Debbah S. Lasaulce and E. Altman, "Methodologies for analyzing equilibria in wireless games," *IEEE Signal Processing Magazine, Special issue on Game Theory for Signal Processing*, vol. 26, no. 5, pp. 41–52, 2009.

Bibliographie

[1] R. Ahlswede. Coloring hypergraphs : A new approach to multi-user source coding, 1. *Journal of combinatorics, information and system sciences*, 4(1), 1979.

[2] R. Ahlswede. Coloring hypergraphs : A new approach to multi-user source coding, 2. *Journal of combinatorics, information and system sciences*, 5(3), 1980.

[3] R. J. Aumann. Survey of repeated games. *Essays in game theory and mathematical economics : in honor of Oskar Morgenstern.*, 1981.

[4] D. Austen-Smith. Strategic transmission of costly information. *Econometrica : Journal of the Econometric Society*, pages 955–963, 1994.

[5] G. Bacci, L. Sanguinetti, M. Luise, and H. Poor. A game-theoretic approach for energy-efficient contention-based synchronization in OFDMA systems. *IEEE Transactions on Signal Processing*, 61(5) :1258–1271, 2013.

[6] T. Başar and S. Yüksel. *Stochastic Networked Control Systems*, volume XVIII. Springer, 2013.

[7] S. Baliga and S. Morris. Co-ordination, spillovers, and cheap talk. *Journal of Economic Theory*, 105(2) :450–468, 2002.

[8] S. Baliga and T. Sjöström. The strategy of manipulating conflict. *The American Economic Review*, 102(6) :2897–2922, 2012.

[9] J. S. Banks and R. L. Calvert. A battle-of-the-sexes game with incomplete information. *Games and Economic Behavior*, 4(3) :347–372, 1992.

[10] E. V. Belmega and S. Lasaulce. Energy-efficient precoding for multiple-antenna terminals. *IEEE Trans. Signal Process.*, 59(1) :329–340, January 2011.

[11] A. Bereyhi, M. Bahrami, M. Mirmohseni, and M. Aref. Empirical coordination in a triangular multiterminal network. In *Proc. of IEEE International Symposium on Information Theory*, pages 2149–2153, 2013.

[12] R. Blasco-Serrano, R. Thobaben, and M. Skoglund. Polar codes for coordination in cascade networks. In *Proc. of International Zurich Seminar on Communications*, pages 55–58, Zurich, Switzerland, March 2012.

[13] M. R. Bloch and J. Kliewer. Strong coordination over a line network. In *Proc. IEEE International Symposium on Information Theory*, pages 2319–2323, Istanbul, Turkey, July 2013.

[14] M. R. Bloch and J. Kliewer. Strong coordination over a three-terminal relay network. accepted to *IEEE Information Theory Workshop*, May 2014.

[15] M. R. Bloch, L. Luzzi, and J. Kliewer. Strong coordination with polar codes. In *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*, pages 565–571. IEEE, 2012.

[16] J. A. Bondy and U. S. R. Murty. *Graph theory with applications*, volume 6. Macmillan London, 1976.

[17] S. P. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[18] G. Caire and S. Shamai. On the achievable throughput of a multiantenna Gaussian broadcast channel. *IEEE Transactions on Information Theory*, 49(7) :1691–1706, July 2003.

[19] Y. Chen. Communication with two-sided asymmetric information. *Available at SSRN 1344818*, 2009.

[20] C. Choudhuri, Y.-H. Kim, and U. Mitra. Capacity-distortion trade-off in channels with state. In *Proc. of 48th Annual Allerton Conference on Communication, Control, and Computing*, pages 1311–1318, 2010.

[21] C. Choudhuri, Y.-H. Kim, and U. Mitra. Causal state amplification. In *Proc. of IEEE International Symposium on Information Theory Proceedings*, pages 2110–2114, Saint Petersburg, Russia, August 2011.

[22] C. Choudhuri and U. Mitra. Action dependent strictly causal state communication. In *Proc. of IEEE International Symposium on Information Theory*, pages 3058–3062, 2012.

[23] A. S. Cohen and A. Lapidoth. The Gaussian watermarking game. *IEEE Transactions on Information Theory*, 48(6) :1639–1667, June 2002.

[24] R. Cooper, D. V. DeJong, R. Forsythe, and T. W. Ross. Communication in coordination games. *The Quarterly Journal of Economics*, pages 739–771, 1992.

[25] T. M. Cover and M. Chiang. Duality between channel capacity and rate distortion with two-sided state information. *Information Theory, IEEE Transactions on*, 48(6) :1629–1638, 2002.

[26] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, 2006.

[27] V. Crawford and J. Sobel. Strategic information transmission. *Econometrica : Journal of the Econometric Society*, pages 1431–1451, 1982.

[28] I. Csiszár. The method of types [information theory]. *Information Theory, IEEE Transactions on*, 44(6) :2505–2523, 1998.

[29] P. Cuff. Distributed channel synthesis. *IEEE Transactions on Information Theory*, 59(11) :7071–7096, 2013.

[30] P. Cuff, H. H. Permuter, and T. M. Cover. Coordination capacity. *IEEE Transactions on Information Theory*, 56(9) :4181–4206, 2010.

[31] P. Cuff and C. Schieler. Hybrid codes needed for coordination over the point-to-point channel. In *Proc. of 49th Annual Allerton Conference on Communication, Control, and Computing*, pages 235–239, 2011.

[32] P. Cuff and L. Zhao. Coordination using implicit communication. In *Information Theory Workshop (ITW), 2011 IEEE*, pages 467–471. IEEE, 2011.

[33] G. Debreu. A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences of the United States of America*, 38(10) :886, 1952.

[34] S. Deilami, A. S. Masoum, P. S. Moses, and M. A. S. Masoum. Real-Time Coordination of Plug-In Electric Vehicle Charging in Smart Grids to Minimize Power Losses and Improve Voltage Profile. *IEEE Trans. Smart Grid*, 2(3) :456–467, 2011.

[35] M. Dyer, A. Frieze, and R. Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *Journal of the ACM (JACM)*, 38(1) :1–17, 1991.

[36] M. Emelianenko, L. Ju, and A. Rand. Nondegeneracy and Weak Global Convergence of the Lloyd Algorithm in Rd. *SIAM J. Numerical Analysis*, 46(3) :1423–1441, 2008.

[37] L. Fisher. *Rock, Paper, Scissors : Game Theory in Everyday Life*. Basic Books, 2008.

[38] F. Forges and F. Koessler. Long persuasion games. *Journal of Economic Theory*, 143(1) :1–35, 2008.

[39] B. Fourestié. Communication switching method, access point, network controller and associated computer programs. Patent wo 2008/081137, France Telecom, Dec. 2007.

[40] A. Gamal and Y. Kim. *Network Information Theory*. Cambridge University Press, 2011.

[41] C. Ge, F. Ma, and J. Zhang. A fast and practical method to estimate volumes of convex polytopes. *arXiv preprint arXiv :1401.0120*, 2013.

[42] S. Gel'Fand and M. S. Pinsker. Coding for channel with random parameters. *Probl. Contr. Inform. Theory*, 9(1) :19–31, 1980.

[43] A. Gjendemsj, D. Gesbert, G. E. Oien, and S. G. Kiani. Binary power control for sum rate maximization over multiple interfering links. *IEEE Transactions on Wireless Communications*, 7(8) :3164–3173, 2008.

[44] D. Goodman and N. Mandayam. Power control for wireless data. *IEEE Personal Communications*, 7(2) :45–54, April 2000.

[45] S. Gordon. On infinite cheap talk equilibria. *manuscript, University of Montreal*, 2010.

[46] O. Gossner, P. Hernandez, and A. Neyman. Optimal use of communication resources. *Econometrica*, 74(6) :1603–1636, 2006.

[47] P. Grover. *Actions can speak more clearly than words*. PhD thesis, Univ. of California at Berkeley, Jan. 2011.

[48] F. Haddadpour, M. H. Yassaee, A. Gohari, and M. R. Aref. Coordination via a relay. In *Proc of IEEE International Symposium on Information Theory*, pages 3048–3052, Boston, MA, July 2012.

[49] T. Han and S. Verdú. Approximation theory of output statistics. *IEEE Trans. Inf. Theory*, 39(3) :752–772, May 1993.

[50] S. Haykin. Cognitive radio : brain-empowered wireless communications. *IEEE Journal on Selected Areas in Communications*, 23(2) :201–220, 2005.

[51] IEC. *Loading guide for Oil-Immersed Power Transformers*. Number 60354. IEC, 2011.12.02.

[52] G. Jäger, L. P. Metzger, and F. Riedel. Voronoi languages : Equilibria in cheap-talk games with high-dimensional types and few signals. *Games and economic behavior*, 73(2) :517–537, 2011.

[53] R. Kannan, L. Lovász, and M. Simonovits. Random walks and an $o*(n^5)$ volume algorithm for convex bodies. *Random structures and algorithms*, 11(1) :1–50, 1997.

[54] V. Kavitha, E. Altman, R. El-Azouzi, and R. Sundaresan. Opportunistic scheduling in cellular systems in the presence of noncooperative mobiles. *Information Theory, IEEE Transactions on*, 58(3) :1757–1773, 2012.

[55] Y.-H. Kim, A. Sutivong, and T. Cover. State amplification. *IEEE Transactions on Information Theory*, 54(5) :1850–1859, 2008.

[56] F. Koessler and F. Forges. Transmission stratégique de l'information et certification. *Annales d'Economie et de Statistique*, pages 3–61, 2008.

[57] G. Kramer and S. A. Savari. Communicating probability distributions. *IEEE Transactions on Information Theory*, 53(2) :518–525, February 2007.

[58] S. Lasaulce and H. Tembine. *Game Theory and Learning for Wireless Networks : Fundamentals and Applications*. Academic Press, 2011.

[59] J. B. Lasserre. An analytical expression and an algorithm for the volume of a convex polyhedron in $\mathbb{R}^n$. *Journal of optimization theory and applications*, 39(3) :363–377, 1983.

[60] M. Le Treust and S. Lasaulce. Transforming monitoring structures with resilient encoders—application to repeated games. *Dynamic Games and Applications*, 3(1) :38–67, 2013.

[61] M. Le Treust, A. Zaidi, and S. Lasaulce. An achievable rate region for the broadcast wiretap channel with asymmetric side information. In *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pages 68–75. IEEE, 2011.

[62] Q. Li, D. Gesbert, and N. Gresset. Joint precoding over a master-slave coordination link. In *Proc. of the IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Florence, Italy, May 2014.

[63] V. Marano, S. Onori, Y. Guezennec, G. Rizzoni, and N. Madella. Lithium-ion batteries life estimation for plug-in hybrid electric vehicles. In *Vehicle Power and Propulsion Conference, 2009. VPPC '09. IEEE*, pages 536–543, 2009.

[64] M. Maschler, S. Zamir, and E. Solan. *Game Theory*. Cambridge University Press, 2013.

[65] S. A. Matthews and A. Postlewaite. *On modeling cheap talk in Bayesian games*. Springer, 1994.

[66] N. Merhav and S. Shamai. On joint source-channel coding for the wyner-ziv source and the gel'fand-pinsker channel. *Information Theory, IEEE Transactions on*, 49(11) :2844–2855, 2003.

[67] F. Mériaux, M. L. Treust, S. Lasaulce, and M. Kieffer. A stochastic game formulation of energy-efficient power control : Equilibrium utilities and practical strategies. *CoRR*, abs/1107.4258, 2011.

[68] F. Mériaux, S. Valentin, S. Lasaulce, and M. Kieffer. An energy-efficient power allocation game with selfish channel state reporting in cellular networks. In *Performance Evaluation Methodologies and Tools (VALUETOOLS), 2012 6th International Conference on*, pages 290–297, 2012.

[69] P. Mertikopoulos, E. V. Belmega, A. L. Moustakas, and S. Lasaulce. Distributed learning policies for power allocation in multiple access channels. *IEEE Journal on Selected Areas in Communications*, 30(1) :96–106, January 2012.

[70] F. Meshkati, H. Poor, and S. Schwartz. Energy-efficient resource allocation in wireless networks : An overview of game theoretic approaches. *IEEE Signal Processing Magazine*, 58 :58–68, May 2007.

[71] D. Monderer and L. Shapley. Potential games. *Games and Economic Behavior*, 14(1) :124 – 143, 1996.

[72] J. F. Nash et al. Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, 36(1) :48–49, 1950.

[73] M. Rabin. A model of pre-game communication. *Journal of Economic Theory*, 63(2) :370–391, 1994.

[74] J. Renault and T. Tomala. Repeated proximity games. *International Journal of Game Theory*, 27(4) :539–559, 1998.

[75] A. Rubinstein. Equilibrium in supergames with the overtaking criterion. *Journal of Economic Theory*, 21(1) :1–9, 1979.

[76] M. D. S. Lasaulce and E. Altman. Methodologies for analyzing equilibria in wireless games. *IEEE Signal Processing Magazine, Special issue on Game Theory for Signal Processing*, 26(5) :41–52, 2009.

[77] V. Santos. Alternating-announcements cheap talk. *Journal of Economic Behavior & Organization*, 42(3) :405–416, 2000.

[78] G. Scutari, D. Palomar, and S. Barbarossa. The mimo iterative waterfilling algorithm. *IEEE Transactions on Signal Processing*, 57(5) :1917–1935, May 2009.

[79] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27 :379–423, 1948.

[80] C. E. Shannon et al. Two-way communication channels. In *Proc. 4th Berkeley Symp. Math. Stat. Prob*, volume 1, pages 611–644. USA, 1961.

[81] D. Slepian and J. K. Wolf. Noiseless coding of correlated information sources. *Information Theory, IEEE Transactions on*, 19(4) :471–480, 1973.

[82] J. Sobel. Signaling games. *Encyclopedia of Complexity and System Science, Springer, forthcoming*, 2007.

[83] M. L. Treust and S. Lasaulce. A repeated game formulation of energy-efficient decentralized power control. *IEEE Trans. on Wireless Commun.*, 9(9) :2860–2869, September 2010.

[84] E. Tuncel. Slepian-wolf coding over broadcast channels. *Information Theory, IEEE Transactions on*, 52(4) :1469–1482, 2006.

[85] J. Von Neumann and O. Morgenstern. *Theory of games and economic behavior (2d rev)*. Princeton University Press, 1947.

[86] X. Wu. On convergence of Lloyd's method I. *IEEE Transactions on Information Theory*, 38(1) :171–174, 1992.

[87] W. Yu, G. Ginis, and J. M. Cioffi. Distributed multiuser power control for digital subscriber lines. *IEEE Journal on Selected Areas in Communications*, 20(5) :1105–1115, Jun. 2002.

[88] A. Zappone, S. Buzzi, and E. Jorswieck. Energy-efficient power control and receiver design in relay-assisted DS/CDMA wireless networks via game theory. *IEEE Comm. Lett.*, 15(7) :701–703, July 2011.