HAL open science

Development and Application of Information Theoretical Bounds to Certain Class of Coordination Problems

Achal Agrawal

▶ To cite this version:

Achal Agrawal. Development and Application of Information Theoretical Bounds to Certain Class of Coordination Problems. Signal and Image processing. Université Paris Saclay (COmUE), 2016. English. NNT : 2016SACLS170 . tel-01494660

HAL Id: tel-01494660
https://theses.hal.science/tel-01494660

Submitted on 23 Mar 2017

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

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

NNT - 2016SACLS170

THESE DE DOCTORAT
DE
L'UNIVERSITE PARIS-SACLAY
PREPAREE A
L'UNIVERSITE PARIS SUD

ECOLE DOCTORALE N° 580
Sciences et technologies de l'information et de la communication

Spécialité de doctorat : Réseaux, Information et Communications

Par

M. Achal Agrawal

Development and Application of Information Theoretical Bounds to Certain Class of Coordination Problems

Thèse presentée et soutenue à Gif-Sur-Yvette, le 14 Juin 2016 :

Composition du Jury :

M. Chahed, Tijani	Professeur, Telecom SudParis	Président
M. Gesbert, David	Professeur, EURECOM	Rapporteur
M. Yezekael, Hayel	Maître de Conférences, Université d'Avignon	Rapporteur
M. Assaad, Mohamad	Professeur Adjoint, CentraleSupelec	Examinateur
M. Lasaulce, Samson	Directeur de Recherche, CNRS	Directeur de thèse

List of Figures

Acknowledgment

First and foremost, I would like to sincerely thank my thesis director Prof. Samson Lasaulce for his unflinching support and belief in me. From my experience as representative of the PhD students of L2S, I know that such support and encouragement is indeed rare, and I was extremely lucky in this respect.

I would also like to thank project LIMICOS - ANR-12-BS03-0005 for financing the major part of my thesis, as well as Orange Labs, Issy-les-Moullineux for financing the remainder of my thesis.

It was perhaps no co-incidence that the team headed by Samson had very good synergy and interactions. Vineeth Varma, a former classmate of mine, was my housemate and my mentor for most part of my PhD. I am also very grateful of Benjamin Larrousse, whose work paved the path for the works presented in this thesis, and for the numerous collaborations undertaken with him. Together with Olivier Beaude, he also helped keep up the ambiance in the office, and his presence was sorely missed in the office. I would also like to acknowledge the contribution of Olivier Beaude in trying to motivate the team and being a model for the rest of us to follow in terms of efficiency and hard work. More recently, thanks to Chao Zhang, and to a lesser extent Sara Berri, I managed to pass through the toughest times for a PhD student, the end! Florian Danard stayed only briefly in the team, but he left a lasting impression on me with his humility, given his unquestionable mathematical strengths.

I would also like to thank Pierre Gerold, Elsa Dupraz, Franois Meriaux, Chengfang Ren, for various interesting discussions we had over lunches and coffee breaks. My smoking friends Thang, Mircea, Rajesh and Angelo helped me keep it real with some freewheeling chats. My mexican friends, Pablo and Rafa in particular were great ping-pong buddies, and really cool people.

On a more personal note, I am deeply endebted to my close friends Vianney Costemalle, Mang Chetra, Bhanukiran, Alfredo Alarcon, Chloe Massin for helping me pull through the PhD and for bearing with my erraticness.

A very big credit of my finishing my thesis goes to my love, Diana Markozashvilli, who made me believe in myself and made me take up my responsibilities when the going got tough.

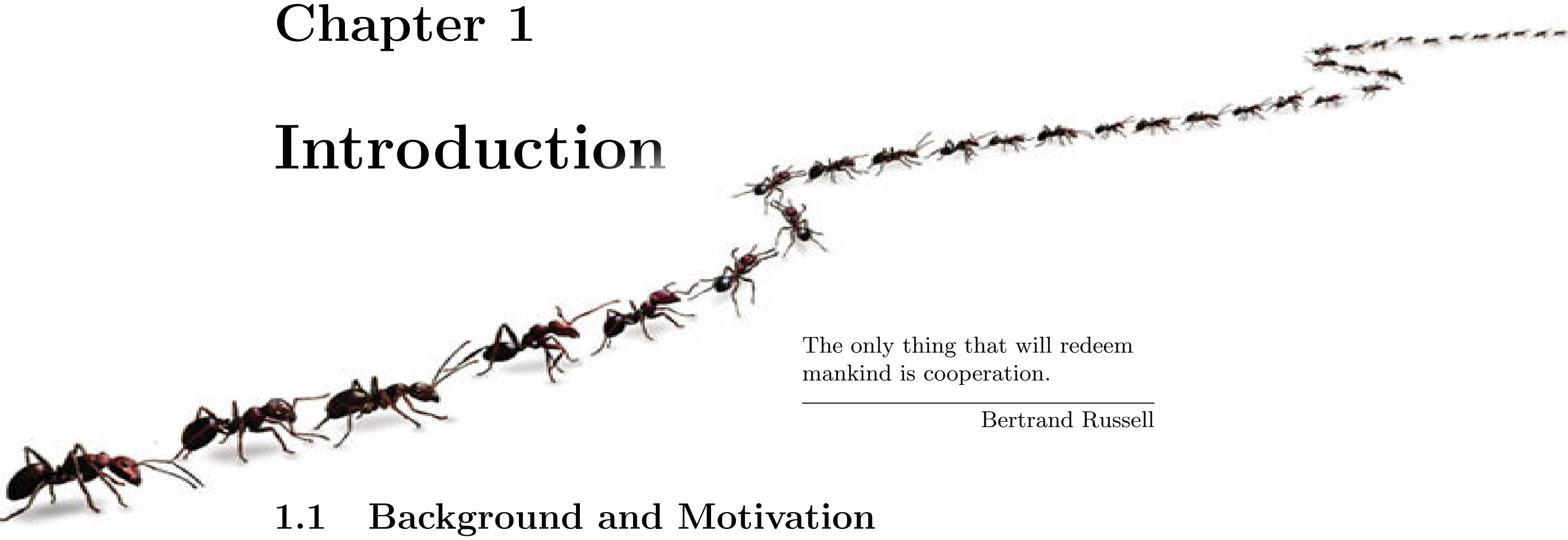
Lastly, but not in the very least, I would like to thank my parents for supporting my decision to do research, against their wishes.

Contents

I thought the grass was greener on the other side, but...
I was only chasing the pot at the end of the rainbow.
Now I know the straight dope.

Chapter 1

Introduction



The only thing that will redeem mankind is cooperation.

Bertrand Russell

1.1 Background and Motivation

In recent times, our computing and sensing capabilities have grown manyfold thanks to accelerated innovation rhythm worldwide. This has led to a proliferation of digital devices. In addition, most devices are increasingly getting connected and are capable of receiving, and in some cases sending information. While this connectivity is currently exploited mostly for providing added functions to devices, various new opportunities exist to exploit this new-found ubiquity of information generated for helping devices coordinate better amongst themselves for certain *common* objectives or goals. A few examples of such common objectives, include, but are not limited to: coordination amongst self driving cars, coordination of drones to perform a common task, coordination between shops for logistics etc. These payoffs thus depend on the decision of every individual coordinating *agent*.

Usually, the decisions need to be coordinated taking into account certain externalities which are random. If this were not the case, there would be no information available w.r.t. which the agents would need to coordinate. Also their strategies would be stationary based solely on the estimated statistics of the externalities. Due to this randomness in the *nature state* and its effect on the payoff, it is more interesting to evaluate schemes which maximize expected payoff w.r.t. the nature state. Thus the objective of the optimization problem is that of finding coordination schemes for maximizing expected payoffs. While, in general, the realizations of this nature state might be correlated, we shall consider the case where they are independent and identically distributed (iid). The performance achieved for the i.i.d. case can only be bettered if there actually were correlations which were taken into account.

While centralized solutions to optimization problems are attractive due to better promised performances, they are highly impractical since functional optimization often suffers with prohibitive complexity. Its impracticality also stems from the extensive backhaul required to relay the relevant information upto the central decision-making entity and then relaying back the decision chosen to the agents. Again, this only delays the time required to make the decision. In rapidly changing environments, this time delay could render the whole exercise futile. Moreover, such solutions are not robust to noise and vulnerable to total network failures due to propagation of errors.

A more practical approach is to consider *decentralized decision-making*, with the understanding that the agents co-operate amongst themselves (with or without explicitly communicating) and making informed decisions. The co-operation is necessary as otherwise the solutions might suffer from the problem of operating at Nash equilibria which are generally not socially optimal [4]. Eventhough nash equilibria result from certain desirable assumptions about decisions taken in a non-cooperative scheme; one does not need a Karl Marx to realize that better performance could be achieved through co-operation at the expense of certain agents.

Nature has provided a good example of what decentralized cooperation schemes can achieve in the form of ants. Despite very limited computation and sensing capabilities, their resourcefulness in sourcing food and creating complex structures is well documented. Indeed, many mathematically similar problems in traffic control, routing protocols have been explored using algorithms inspired by ants.

For illustrative purposes, consider the mechanism ants use in sourcing food. Ants manage to find the shortest paths from food source to colony without the ability to either sense the food at a distance or explicitly communicate its location to other ants. Ants lay a pheromone trail behind when they walk, which helps them retrace their path back to the colony once food is found. The pheromones however serve another purpose, that of communication with other ants. When an ant, after having found food, returns to the colony, it reinforces the path taken with more pheromones. Ants simply follow trails with stronger pheromone scent and over time this algorithm converges to solutions close to optimal. Notice however that the communication in this case is not explicit. The communication happens implicitly between ants by an ant stumbling across the pheromone trail of other ants. Also, each ant acts only on locally available information: the strength of the pheromone scent. Shorter paths are chosen in a natural manner as the frequency of an ant doing to-and-fro passing through a point in a shorter path is higher.

The decentralization of decision-making leads to the information available to different decision makers being different. This difference could manifest itself due to a plethora of reasons:

- Availability of only local information - In general, devices are only connected to other devices and sensors in a local neighbourhood. Even if they are connected further, rapidly changing nature state could force agents to adapt faster, and thus with more localized information.

- Hierarchy of information - Some connected devices might act as nodes or aggregators for a small group, having more information. Some nodes might even have information about future realizations or forecasts of the nature state with respect to which the devices are coordinating.

- Noisy channels - Even when the information transmitted to different devices is the same, each device will observe a noisy version of the signal which will be different for each device.

In general, these three parameters define the *information structure* of a given coordination problem. Indeed, the performance that can be achieved by coordination is constrained by what information each agent can use to make his decisions and coordinate. Thus, the payoffs that are achievable are characterized by the information structure. Such optimization problems are known as *team-decision problems* and have been studied extensively since Marschak published his seminal work [5]. Together with [6], it formed the basis for team-decision theory. Since then, team-decision problems have been considered with various information structures, and is still an active field of research.

While the ants have devised a decentralized scheme, it might or might not make optimal use of the information available to each ant. The optimality of the scheme depends on information theoretical limits of implicit communication i.e. communication achieved while acting towards an objective. We use certain information theoretical results proved recently [7] [8], [9] which provide precisely the performance limits of coordination problems for certain information structures. Eventhough these results are fairly new and make some assumptions about the information structure, they provide good frameworks for tackling certain interesting applications in various domains such as Telecommunications, Automatic Control, and Smart Grids to name a few.

We divide this thesis into two parts based on two different information structures considered for applications. In the first part, we consider a non-causal information structure with a hierarchy of information where one agent knows the future realizations of the nature state in a non causal manner. Other agents only observe the actions chosen by the informed agent and make their decisions accordingly. Second part explores a causal information structure, i.e. all agents can only observe past or present, but not the future. Two scenarios are considered for the causal case; where agents observe an image of the nature state, or where agents also observe the actions of other agents in a causal manner. The former is a special case of the latter.

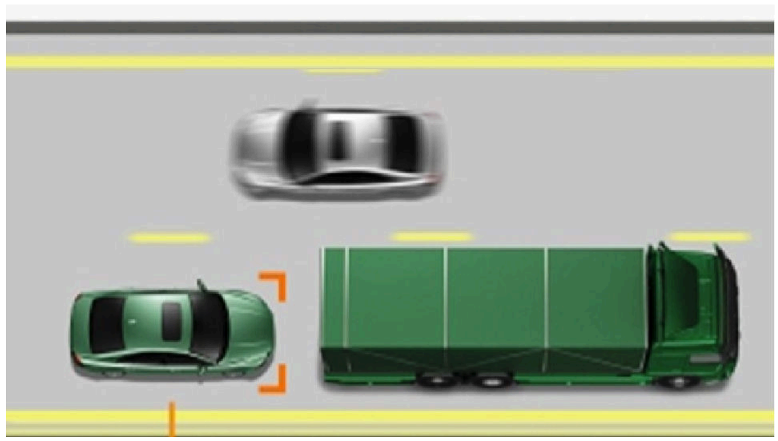


Figure 1.1: Truck and car co-ordinating to stay on course

1.2 Coordination problems with a non-causal information structure

Inspired from [7], we considered 2-Agent coordination problems with non causal information structure. We assume that one agent has complete and noncausal knowledge of the sequence of i.i.d. realizations of the nature state X_0 for the entire time period of optimization. The informed agent can exchange his knowledge with the other agent only through his actions. While the non causal assumption might seem unrealistic, one can find many scenarios where it holds true.

Consider the following example of a truck and a car coordinating with the goal of safe navigation over a non-straight road. In this example, the truck, having superior visual information, needs to communicate information about the path to the car. This can be achieved through many ways. A simple solution is to have a communication channel between the two for this purpose. However, imagine that no such channel exists, and the car is reliant solely on the movements of the truck to predict the path ahead. The dual role of the movements of truck, to stay on course as well as signal the path ahead to the car, is the defining characteristic of implicit communication. For further discussion and interesting examples of implicit communication, the author refers to the introduction of [10] .

We consider this example for many instructive reasons. Firstly, it is a practical scenario being tested by Volvo for creating road trains as a solution for self-driving traffic [11] . Secondly, it naturally has the non-causal information structure considered in [12] [7] as the truck can see the path ahead whereas the car behind has no direct visuals of the path ahead and relies solely on signalling by the truck for path prediction. Thirdly, this example is closely related to the famous Witsenhausen counterexample in control theory. We shall explore the last point in detail later.

Indeed, similar information structure has been treated before. Notably,The case of perfect observation by agents is treated in [12] while the generalization to noisy observations by agents is conducted in [7]. 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 states realizations at all induces no limiting performance loss [13]. Reference [7] 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 contributions of our work is precisely to study this general problem in detail by applying it to solve a real optimization problem in resource allocation.

We distinguish two cases for the non-causal information structures: 1) Discrete Case - where all the intervening variables, system state X_0, and the actions chosen by the users $X_i, i \in \{1, 2\}$ are from a discrete alphabet, and 2) Continuous case - where those alphabets are continuous.

1.2.1 Discrete Case: Solving the optimization problem

In [7] , they characterize the performance limit of coordination problems with the aforementioned non causal information structure. In particular, they generalize the result found in [12] which only treated the case where the second agent has a perfect observation of the first agent's actions. However, both the references do not solve the optimization problem found with information theoretic constraints. We consider the optimization problem for the case of perfect observations by agent 2, and show certain properties of the solutions to this optimization problem under some reasonable assumptions about the payoffs.

The results thus obtained are 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 setting, the nature state is given by the global channel state. The informed transmitter uses a sequence of power vectors as a code which conveys information about the channel to the other transmitter.

We consider the discrete case, i.e. discrete power levels, and channel coefficients. While majority of the literature on power allocation optimization over multi-band channels (starting with the pioneering work[14]) consider In contrast with the vast majority of related works on distributed power allocation over , 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., [15][16][17][18]).

1.2.2 Continuous Case: Generalizing the Information Constraint

However, the variables could in general take values in a continuous alphabet. To this effect, we generalized the information theoretic bounds in [7] to the case when the alphabets are continuous. This generalization is not so straightforward as one needs to take care in redefining the information theoretical notions of mutual information. We consider 2 scenarios for the continuous case: the second agent causally observing the actions of the informed agent, or it observes the realization of nature state too in a causal manner, in addition to the actions of the informed agent. We provide the Information Constraints for both the scenarios for the continuous case. We also treat the special case of all the variables being distributed normally for both the scenarios considered. Also, we treat the case of perfect observation and explicit the information constraint if the variables follow log-normal distribution and burr's distribution.

Another contribution for the continuous case is to consider the Witsenhausen cost function [19] as a common cost function to be minimized by the $2-$agent team under the two mentioned scenarios in terms of information structures; this establishes for the first time a link between [12][20][21] and [19]. Indeed, although the Witsenhausen problem can be seen as a one-shot coordination problem, whereas we consider a long-term coordination problem here, the idea of joint control-communication strategies is present in both formulations. As it will be seen, characterizing the feasible performance of the long-term coordination problems amounts to determining a certain information constraint. Although information constraints normally appear when (large) sequences intervene, it has been noted that one-shot problems closely related to the Witsenhausen problem have been solved by introducing an information constraint; this is the case, for instance, for the Gaussian Test Channel (GTC) [22].

We treat the original Witsenhausen counterexample using the theory developed in part II as the information structure therein is more adapted to it.

1.3 Coordination problems with a causal information structure

The non-causal information structure however is not applicable to scenarios where only causal observations might be available. In general however, its reasonable to assume that future realizations of the nature state are not known to any agent, even if their statistics are known. In [8], they characterize the achievable payoffs for coordination with such causal information structure. In particular, they show that the achievable payoffs are characterized by conditional probability distributions for all agents (representing the decisions made). Each conditional probability distribution is shown to be independent of each other. This independence comes directly due to the information structure as all agents are supposed to have an image of the nature state through independent channels. In other words, they do not observe each others' actions. Note that it is this difference in information structure which makes the information constraints found in the non-causal case irrelevant.

We use the independence of decisions to construct a distributed algorithm based on sequential global best response dynamics (BRD) which aims at optimizing the payoff for the given structure. While we provide no proof for optimality, we show via numerous applications that it indeed provides a good framework for tackling practical coordination problems in various settings.

The benefits of our framework, when compared to existing literature on decentralized decision making for the aforementioned applications, essentially boils down to a simple reason; cooperation. Most of the current decentralized solutions evaluate non-co-operative games, and provide strategies to attain Nash Equilibria. While this approach is robust, by virtue of its assumed selfishness of the agents while making their decisions, it clearly does not aim to provide optimal performances assuming cooperation.

The strategies found using our decentralized algorithm has the following salient features which make it more efficient, as well as robust and implementable.

- Cooperative strategies - Since we solve the optimization problem with the goal of optimizing common objective, typically sum-utilities of the individual agents, we easily beat non-cooperative schemes. Moreover, since we use information bound optimal payoffs, we achieve strategies which achieve close to optimal performance for the considered information structure.

- Decentralized decision making - Each agent makes its decision based on local information available to him taking into account the global network statistics. Thus the decentralization of decision is based on decentralized information structure. In particular, we can evaluate the payoff corresponding to different partial information of the nature state available to the coordinating agents.

- Robustness to noise - Indeed, in practical cases the feedback has a noise associated with it. In the algorithms proposed, we take into account the noise statistics for generating the best responses, creating robust functions w.r.t. feedback noise.

- Offline optimization - Agents can find their decision functions individually provided they have the global network statistics, which are easily received through network backhauls. They use an iterative method to find functions maximising expected utility given the probability distributions of network parameters. Thus they can infer the best response of each agent sequentially and converge to its optimal best response.

We identified and modelled coordination problems in two major technological challenges: 1) Power optimization schemes proposed could help improve technological standards for 5G technologies by making them more power efficient and reducing interference caused by independent terminals. 2) Communication within the electricity distribution network can be used, along with neighbourhood consumption forecasts, to even out the consumption over time, thus reducing neighbourhood transformer ageing and joules losses related to transmission costs. In addition to these applications, we treat the famous Witsenhausen

Counterexample [19] which has a slightly different information structure, where an agent can also observe the actions of another agent in a causal manner.

1.3.1 Power optimization for Wireless Communications with partial CSI

Most power optimization schemes for Wireless communications consider non-cooperative strategies and employ game theory to provide guarantees on performance and convergence. While these schemes are robust, the performance achieved by them are not socially optimal as they converge to Nash Equilibria. We use the characterization explicited in [8] to create co-ordination schemes which are sub-optimal, but still perform better than the solutions proposed till now. Due to the generality of the theory developed, we provide a framework which can be used to tackle different utilities sum-rate, sum-energy, sum-throughput under different settings like parallel interference channel and multiple access channel.

We pay greater attention to sum-energy as it is the major focus of wireless communications in recent times. We show that a combination of thresholding and channel inversion strategies provide good performances by mitigating the interference in case the communication channels are bad.

1.3.2 Power consumption scheduling in Smart Grids with uncertainty

The main aim was to make a step further towards knowing how an electrical appliance should exploit the available information to schedule its power consumption. This information corresponds here to an imperfect forecast of the non-controllable (exogenous) load or electricity price. Reaching this goal led us to three key results which can be used for other settings which involve multiple agents with partial information

- In terms of modeling, we exploit the principal component analysis to approximate the exogenous load and show its full relevance
- Under some reasonable but improvable assumptions, this work provides a full characterization of the set of feasible payoffs which can be reached by a set of appliances having partial information
- A distributed algorithm is provided to compute good power consumption scheduling functions. These results are exploited in the numerical analysis, which provides several new insights into the power consumption scheduling problem.

We provide first results for the standard cost functions, transformer aging in particular, where we compare our method with iterative water filling algorithm (IWFA). We test our proposed algorithm on real data and show that it is more robust with respect to noise in the signals received. We also observe that our proposed method becomes even more relevant when the proportion of appliances with smart counters increase.

1.3.3 Witsenhausen Counterexample

The famous Witsenhausen counterexample has been a key toy problem for dual objectives of control and communication. Its difficulty can be gauged by the fact that the search for optimal control functions for the problem has been unresolved since 1968 . We generalize the theorem provided by [8] to the case where agents can observe the action of others in a causal manner to get some insights into this problem.

We propose an iterative approach to solve the optimization problem since finding global optimum has prohibitive complexity. This simplification leads us to an algorithm similar to [23], albeit with certain small differences in the modelization of observation and action spaces. The differences in quantization can be justified due to [24], which shows that finite and uniform quantized alphabets for observations give strategies with are *epsilon-optimal*. We thus obtain functions which are very close to the best solutions known to date (difference of 0.3%), and have a smaller complexity than [23].

1.4 Thesis Outline

This thesis consists of 2 parts. The first part comprises of chapter 2 and chapter 3 and deals extensively with the proposed non-causal information structure. The second part includes chapter 4 and chapter 5 and treats the causal information structure.

In **Chapter 2**, we introduce the non-causal information structure as well as recall the information theoretical results which we attempt to further develop. These results were shown for the discrete case, i.e. all action alphabets as well as alphabets representing the nature state being discrete. We go further and solve an optimization problem stated by the articles providing the theoretical background, and provide first insights into the structure of the possible solutions.

Chapter 3 generalizes the aforementioned results for the continuous case. While traditionally, the generalization from discrete information theoretical results to continuous ones is assumed, the generalization is not so straightforward as certain concepts need to be redefined carefully. As an application, we consider the Witsenhausen cost function, inspired from the Witsenhausen counterexample whose original version we treat in chapter 5.

In the second part of the thesis, we treat the causal information structure. A general framework is developed in **Chapter 4** which recalls the information theoretical characterization. We then propose a decentralized algorithm which exploits the characterization to provide good and practical coordination schemes in general.

We apply the schemes developed to various applications in different domains in **Chapter 5**; namely team-power optimization in wireless networks in 5.1, power consumption scheduling in smart grids in 5.2, and finding decision functions for Witsenhausen Counterexample in **??**. The proposed decision functions outperform the current state of the art simply because of the built-in cooperation while determining them. For Witsenhausen Counterexample, the algorithm proposed coincides with an algorithm already proposed in the literature heuristically.

Concluding remarks are provided in **Chapter 6**.

1.5 Publications

Journal articles

- A. Agrawal, C. Zhang, S. Lasaulce and R. Visoz, "Team Power Optimization using CSI Statistics with Partial Information. Application to Energy Efficient Communications", Transactions on Wireless Communications (TWC), - *to be submitted*

Conference articles

- A. Agrawal, S. Lasaulce, O. Beaude and R. Visoz, "A Framework for Optimal Decentralized Power Control with Partial CSI" , The Fifth International Conference on Communications and Networking, Hammamet, Tunisia, Nov. 2015.

- O. Beaude, A. Agrawal, S. Lasaulce, "A Framework for Computing Power Consumption Scheduling Functions Under Uncertainty", IEEE International Conference on Smart Grid Communications, Miami ,USA Nov. 2015.

- A. Agrawal, F. Danard, B. Larrousse, and S. Lasaulce, "Implicit Coordination in Two-Agent Team Problems with Continuous Action Sets. Application to the Witsenhausen Cost Function", European Conference on Control (ECC), Linz, Austria, June 2015.

- 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.

Part I

Non Causal Information Structure

Chapter 2

Coding through Actions - Discrete Alphabets

> The only reality we truly comprehend is that of our own experience ... The laws of the infinite are extrapolations of our experiences with the finite.
>
> Paul Cohen

2.1 Introduction

In this part, we shall only concentrate on the non-causal information structure briefly introduced in the previous chapter. This information structure has an information hierarchy, with an agent knowing, in a non-causal manner, the realizations of the nature state for the entire time period of coordination. The other agent merely has causal observations of the actions of the 'informed' agent. We only consider the $2-$Agent case in this part as the more general case of many agents is, for now, untreatable. Nonetheless, the 2-Agent case does provide valuable insights into the limits achievable by co-ordination under such an information structure.

We distinguish two different cases: discrete or continuous alphabets for the nature state as well as agents' actions. The information theoretical results of [7] for the discrete case do not generalize to the continuous case in a facile manner. This stems from the difference in interpretation of entropy and mutual information while passing from discrete to continuous alphabets. We prove the corresponding information theoretical bounds for the continuous case in Chapter 3

In this chapter, we discuss the information theoretical bounds for 2-Agent team coordination problems with non-causal information structure due to [7], [9].The central contribution of this chapter 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. As a first step, we restrict our analysis to the case of *perfect observation* by the agents. We use the insights gained to solve a toy power allocation problem in the cognitive radio setting.

The information structure considered in this chapter was first treated by [12]. In that, two agents coordinate over a long period of time, composed of many stages, wherein at each stage they coordinate with the nature state $x_0 \in \mathcal{X}_0$, $|\mathcal{X}_0| < +\infty$ which is assumed i.i.d.. At each stage $t \in \mathbb{T}$, agent $i \in \{1, 2\}$ chooses an action $x_i \in \mathcal{X}_i$, $|\mathcal{X}_i| < +\infty$, which results in the common team payoff $w(x_0, x_1, x_2)$. Additionally, it is assumed that one agent, agent 1, knows beforehand and perfectly all the realizations of the nature state. 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. While in [12] the second agent has a perfect observation of the actions of agent 1, [20] generalizes the results to noisy observations. 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 nature state's realizations at all induces no limiting performance loss [13].

The performance analysis of this problem leads to deriving an information-theoretic constraint. Reference [20] also states an optimization problem which essentially amounts to maximizing the long term payoff function under some constraints, with the information-theoretic constraint being one of them. However, the optimization problem is not analyzed, and this is precisely the main aim of this chapter. As an example of the application

The application of interest in this chapter 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. 2.1. We assume the set of power allocation vectors at a transmitter to be discrete and finite (namely, $|\mathcal{X}_i| < +\infty$) instead of being continuous. This choice is motivated by well-known results in information theory [25] 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. 2.1, 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 chapter is organized as follows. We recall the information theoretical results of [7] in 2.2. In Sec. 2.3, we introduce and solve the general optimization problem of interest. In Sec. 2.4, we apply the general result of Sec. 2.3 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.

2.2 Information Theoretical Bounds

Before we state the general optimization problem with the information theoretical constraints due to the information structure, we would like to recall some results from [7]. Firstly, we restrict the strategies possible for both agents at each time stage t by restricting the variables they can depend on. This also fixes the information structure under consideration. Also, as stated before, we only treat the case of perfect observation by the second agent.

$$\begin{cases} \sigma_t & : & \mathcal{X}_0^T & \rightarrow & \mathcal{X}_1 \\ \tau_t & : & \mathcal{X}_1^{t-1} & \rightarrow & \mathcal{X}_2 \end{cases} \tag{2.1}$$

where σ_t is the strategy for Agent 1, which depends on the knowledge of $\mathcal{X}_0^T$ which denotes the realizations of Nature state $X_0 \in \mathcal{X}_0$ for all the stages T. Similarly, τ_t is the strategy of Agent 2 and depends on the actions of agent 1 $X_1 \in \mathcal{X}_1$ upto, but not including the stage t. The realization of the nature state at stage t, denoted by $x_{0,t}$, as well as the actions chosen by the agents at that stage $x_{1,t}, x_{2,t}$, together result in the payoff at stage t being $w_t(x_{0,t}, x_{1,t}, x_{2,t})$. The aim of the agents is to maximize the payoff on an average over many stages. The average payoff can be calculated as $\overline{w} = \frac{1}{T}\sum_{t=1}^{T} w_t(x_{0,t}, x_{1,t}, x_{2,t})$. Henceforth, we shall analyze the asymptotic case of $T \rightarrow \infty$, which is the limiting case corresponding to the maximum achievable payoff under the given information structure. Since we consider the asymptotic case, the empirical distribution of the random variables $q(X_0X_1X_2, t)$ converges in probability to a time averaged distribution $\overline{q}(X_0, X_1, X_2)$ which characterizes the expected payoff. This is encapsulated in the following definition:

Definition 1 (Implementability)**.** *Let 2.1 be the assumed information structure. The probability density function $\overline{q}(x_0, x_1, x_2)$ is implementable if there exists a pair of control strategies (σ_t, τ_t) such that as $T \rightarrow +\infty$, we have for all $(x_0, x_1, x_2) \in \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2$,*

$$\frac{1}{T}\sum_{t=1}^{T} q_{X_0X_1X_2,t}(x_0, x_1, x_2, t) \rightarrow \overline{q}(x_0, x_1, x_2) \tag{2.2}$$

where $q_{X_0X_1X_2,t} = q_{X_1,X_2|X_0,t} \times q_{X_0}$ is the joint distribution induced by (σ_t, τ_t) at stage t.

Note that since the expectation value of the payoff is a linear operator with respect to the distribution $\overline{q}$, the time averaged expected payoff $\overline{w}$ is reachable if and only if the corresponding distribution $\overline{q}$ is implementable. Also, the factorization of the empirical distribution q can be easily understood, as the random variable X_0 is not controlled by agents and is an extraneous variable. Therefore this factorization is natural. Dropping the index for stage t, we get the expected payoff at a particular stage to be :

$$\begin{aligned}\mathbb{E}[w] &= \sum_{(x_0,x_1,x_2)\in\mathcal{X}_0\times\mathcal{X}_1\times\mathcal{X}_2} q(x_0,x_1,x_2)w(x_0,x_1,x_2)\\ &= \sum_{(x_0,x_1,x_2)\in\mathcal{X}_0\times\mathcal{X}_1\times\mathcal{X}_2} q_{X_0}(x_0)q(x_1,x_2|x_0)w(x_0,x_1,x_2)\end{aligned}$$

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 thus is fixed (referred to as α_i later in this chapter).

However, given the information structure 2.1, which distibutions are implementable, or equivaently, which payoffs are achievable? This is precisely the result of a theorem from [20]

Theorem 1. *[12][20] Let q be a ditribution in $\Delta(\mathcal{X}_0\times\mathcal{X}_1\times\mathcal{X}_2)$ such that $\forall x_0 \in \mathcal{X}_0$, $\sum_{x_1,x_2} q(x_0,x_1,x_2) = q_{X_0}(x_0)$. This distribution q is implementable under the information structure 2.1 if and only if it satisfies the following information constraint:*

$$I_q(X_0;X_2) - H_q(X_1|X_0,X_2) \leq 0 \tag{2.3}$$

where, for any two random variables $(X,Y) \in (X \times Y)$ with joint law $Q(,)$:

- $H_q(X|Y)$ is the conditional entropy of X given Y defined by:

$$H_q(X|Y) = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} q(x,y)\log_2 \frac{q(x,y)}{q_Y(y)} \tag{2.4}$$

  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:

$$H_q(X) = -\sum_{x\in\mathcal{X}} q_X(x)\log_2 q_X(x) \tag{2.5}$$

- $I_q(X;Y)$ denotes the mutual information between X and Y , defined by:

$$I_q(X;Y) = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} q(x,y)\log_2 \frac{q(x,y)}{q_X(x)q_Y(y)} \tag{2.6}$$

Reference [20] 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.

The Information Constraint (2.3) can be re-written as:

$$\begin{aligned} ic(q) &\triangleq I_q(X_0;X_2) - H_q(X_1|X_0,X_2) &(2.7)\\ &= H_q(X_0) + H_q(X_2) - H_q(X_0,X_1,X_2) &(2.8)\end{aligned}$$

2.3 Optimization problem analysis

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 $w = (w_1, w_2, ..., w_n) \in \mathbb{R}^n$ with $n = n_0 n_1 n_2$ 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. 2.1. This choice simplifies the analysis of the optimization problem which is stated next.

Index (i)	X_0	X_1	X_2
1	1	1	1
2	1	1	2
$\vdots$	$\vdots$	$\vdots$	$\vdots$
n_2	1	1	n_2
$n_2 + 1$	1	2	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$2n_2$	1	2	n_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n_2(n_1 - 1) + 1$	1	n_1	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n_1 n_2$	1	n_1	n_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n_1 n_2(n_0 - 1) + 1$	n_0	1	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n_0 n_1 n_2$	n_0	n_1	n_2

Table 2.1: Chosen indexation for the payoff vector w and distribution vector q. Bold lines delineate blocks of size $n_1 n_2$ and each block corresponds to a given value of the random state X_0.

Rewriting the information constraint (2.8) using the chosen indexation, we have:

$$H_q(X_0) = -\sum_{i=1}^{n_0} \left[\Big(\sum_{j=1+(i1)n_1 n_2}^{i n_1 n_2} q_j \Big) \log_2 \Big(\sum_{j=1+(i1)n_1 n_2}^{i n_1 n_2} q_j \Big) \right] \tag{2.9}$$

$$H_q(X_2) = -\sum_{i=1}^{n_2}\left[(\sum_{j=0}^{n_0n_1-1} q_{i+jn_2})\log_2(\sum_{j=0}^{n_0n_1-1} q_{i+jn_2})\right] \tag{2.10}$$

and

$$H_q(X_0, X_1, X_2) = -\sum_{i=1}^{n_0n_1n_2} q_i \log_2 q_i \tag{2.11}$$

The optimization problem of interest consists in 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:

$$\begin{aligned}
\text{minimize} \quad & -\mathbb{E}_q[w] = -\sum_{i=1}^{n_0n_1n_2} q_i w_i \\
\text{s.t.} \quad & -1 + \sum_{i=1}^{n_0n_1n_2} q_i = 0 \\
& -\alpha_i + \sum_{j=1+(i-1)n_1n_2}^{in_1n_2} q_j = 0, \qquad \forall i \in \{1,\dots,n_0\} \\
& -q_i \leq 0,\ \forall i \in \{1,2,\dots,n_0n_1n_2\} \\
& \Bigg\{ -\sum_{i=1}^{n_0}\left[(\sum_{j=1+(i-1)n_1n_2}^{in_1n_2} q_j)\log_2(\sum_{j=1+(i-1)n_1n_2}^{in_1n_2} q_j)\right] \\
& -\sum_{i=1}^{n_2}\left[(\sum_{j=0}^{n_0n_1-1} q_{i+jn_2})\log_2(\sum_{j=0}^{n_0n_1-1} q_{i+jn_2})\right] \\
& + \sum_{i=1}^{n_0n_1n_2} q_i \log_2 q_i \Bigg\} \leq 0
\end{aligned} \tag{2.12}$$

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 which corresponds to (2.3) and has been re-written here as $H_q(X_0) + H_q(X_2) - H_q(X_0, X_1, X_2) \leq 0$, to simplify the analysis.

To solve the optimization problem (2.12) we will apply the Karush Kuhn Tucker (KKT) necessary conditions for optimality [26]. 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 (2.12) and that (2.12) is a convex problem. First, by specializing Lemma 1 in [20] in the case of perfect observation, we know that (2.3) 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 follows from 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* (2.12).

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:

$$
\begin{aligned}
I_q(X_0; X_2) - H_q(X_1|X_0, X_2) &\overset{(a)}{=} 0 - H_q(X_1|X_0, X_2) \\
&\overset{(b)}{=} -H_q(X_1) \\
&\overset{(c)}{<} 0
\end{aligned}
$$

where: (a) and (b) comes from the independence hypothesis between X_0, X_1, and X_2; (c) 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. $\square$

Following the previous considerations, KKT conditions can be applied. The Lagrangian function can be written as:

$$
\begin{aligned}
\mathcal{L}(q, \mu, \mu_0, \lambda, \lambda_{\mathrm{IC}}) = &- \sum_{i=1}^{n_0 n_1 n_2} (w_i q_i + \lambda_i q_i) \\
&+ \mu_0 \left[\sum_{i=1}^{n_0 n_1 n_2} q_i - 1 \right] + \sum_{i=1}^{n_0} \mu_i \left[\sum_{j=1+(i-1)n_1 n_2}^{i n_1 n_2} q_j - \alpha_i \right] \\
&+ \lambda_{\mathrm{IC}} \Bigg\{ - \sum_{i=1}^{n_0} \left[\Big(\sum_{j=1+(i-1)n_1 n_2}^{i n_1 n_2} q_j \Big) \log_2 \Big(\sum_{j=1+(i-1)n_1 n_2}^{i n_1 n_2} q_j \Big) \right] \\
&- \sum_{i=1}^{n_2} \left[\Big(\sum_{j=0}^{n_0 n_1 - 1} q_{i+j n_2} \Big) \log_2 \Big(\sum_{j=0}^{n_0 n_1 - 1} q_{i+j n_2} \Big) \right] \\
&+ \sum_{i=1}^{n_0 n_1 n_2} q_i \log_2 q_i \Bigg\}
\end{aligned}
$$

where $\lambda = (\lambda_1, ..., \lambda_{n_0 n_1 n_2})$, $\mu = (\mu_1, ..., \mu_{n_0 n_1 n_2})$, and IC stands for information-theoretic constraint. KKT conditions follow:

$$\frac{\partial \mathcal{L}}{\partial q_i} = -w_i - \lambda_i + \mu_0 + \left(\sum_{j=1}^{n_0} \mu_j \mathbb{1}_{\{1+n_1 n_2 (j-1) \le i \le j n_1 n_2\}} \right)$$
$$+ \lambda_{\text{IC}} \Bigg[-\sum_{k=1}^{n_0} \mathbb{1}_{\{1+(k-1)n_1 n_2 \le i \le (k) n_1 n_2\}} \log_2 \Big(\sum_{j=1+(k-1)n_1 n_2}^{k n_1 n_2} q_j \Big)$$
$$- \sum_{k=1}^{n_2} \mathbb{1}_{\{i \in \{k, k+n_2, \ldots, k+(n_0 n_1 - 1) n_2\}\}} \log_2 \Big(\sum_{j=0}^{n_0 n_1 - 1} q_{k+j n_2} \Big)$$
$$+ \log_2 q_i - 1 \Bigg] = 0 \qquad \forall\, i \in \{1, 2, \ldots, n_0 n_1 n_2\} \tag{2.13}$$
$$\lambda_i \ge 0 \qquad \forall\, i \in \{1, 2, \ldots, n_0 n_1 n_2\} \tag{2.14}$$
$$\lambda_{\text{IC}} \ge 0 \tag{2.15}$$
$$\lambda_i q_i = 0 \qquad \forall\, i \in \{1, 2, \ldots, n_0 n_1 n_2\} \tag{2.16}$$
$$\lambda_{\text{IC}} i(q) = 0 \tag{2.17}$$

where $\mathbb{1}_.$ is the indicator function and $i(q) =$ is the inequality constraint function associated with the information-theoretic constraint (2.3). 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 (2.12) is such that the information-theoretic constraint is active i.e., $\lambda_{\text{IC}} > 0$.*

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, $\lambda_{\text{IC}} = 0$.

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. 2.1) . 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. 2.1). 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 constraint is active and 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 chapter and thus that $\lambda_{\text{IC}} > 0$. 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 (2.13) can not be satisfied.

- However, if one of the q_i's equals 0, and $q_k = 0$ for all k such that $k[n_2] = i[n_2]$, 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:

- $\lambda_{\text{IC}} > 0$, 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 interesting case that is relevant is:

$$\begin{aligned} \lambda_i &= 0 \qquad \forall\, i \in \{1, 2, \ldots, n_0 n_1 n_2\} \\ \lambda_{\text{IC}} &> 0 \end{aligned}$$

For the latter case, KKT conditions become:

$$\begin{aligned}
&\frac{\partial \mathcal{L}}{\partial q_i} = -w_i + \mu_0 + \left(\sum_{j=1}^{n_0} \mu_j \mathbb{1}_{\{1+n_1 n_2 (j-1) \le i \le j n_1 n_2\}} \right) \\
&+ \lambda_{\text{IC}} \Bigg[- \sum_{k=1}^{n_0} \mathbb{1}_{\{1+(k-1)n_1 n_2 \le i \le (k) n_1 n_2\}} \cdot \\
&\qquad\qquad \log_2 \big(\sum_{j=1+(k-1)n_1 n_2}^{(k) n_1 n_2} q_j \big) \\
&- \sum_{k=1}^{n_2} \mathbb{1}_{\{i \in \{k, k+n_2, \ldots, k+(n_0 n_1 - 1) n_2\}\}} \log_2 \big(\sum_{j=0}^{n_0 n_1 - 1} q_{k+j n_2} \big) \\
&+ \log_2 q_i - 1 \Bigg] = 0 \qquad \forall\, i \in \{1, 2, \ldots, n_0 n_1 n_2\} && (2.18) \\
&\lambda_i \ge 0 \qquad \forall\, i \in \{1, 2, \ldots, n_0 n_1 n_2\} && (2.19) \\
&\lambda_{\text{IC}} \ge 0 && (2.20) \\
&i(q) = 0. && (2.21)
\end{aligned}$$

Now that we have proved some useful results about structure of the optimal solutions of (2.12), 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 (2.12) has a unique solution.*

Proof. We know, by Proposition 2, that $\lambda_{\text{IC}} > 0$ for any optimal solution. It turns out that, if $\lambda_{\text{IC}} > 0$, the Lagrangian of (2.12) is a strictly convex function w.r.t. the vector q. Indeed, as the Lagrangian is the sum of linear functions and a strictly convex function and that the optimization spaces are compact and convex.
It remains to show that $\Phi : Q \mapsto I_Q(X_0; X_2) - H_Q(X_1|X_0, X_2)$ 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 $\lambda_1 \in [0, 1]$, $\lambda_2 = 1 - \lambda_1$, $(Q^1, Q^2) \in \Delta^2(\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2)$ and $Q = \lambda_1 Q^1 + \lambda_2 Q^2$. We have that:

$$\begin{aligned}
H_Q(X_1|X_0X_2) = -\sum_{x_0,x_1,x_2} & \left(\sum_{i=1}^{2} \lambda_i Q^i(x_0, x_1, x_2)\right). \\
& \log\left[\frac{\sum_{i=1}^{2} \lambda_i Q^i(x_0, x_1, x_2)}{\sum_{i=1}^{2} \lambda_i Q^i_{X_2}(x_2)}\right] \\
> -\sum_{x_0,x_1,x_2} \sum_{i=1}^{2} & \lambda_i Q^i(x_0, x_1, x_2) \log\left[\frac{\lambda_i Q^i(x_0, x_1, x_2)}{\lambda_i Q^i_{X_2}(x_2)}\right] \\
= -\sum_{i=1}^{2} \lambda_i & \sum_{x_0,x_1,x_2} Q^i(x_0, x_1, x_2) \log\left[\frac{Q^i(x_0, x_1, x_2)}{Q^i_{X_2}(x_2)}\right] \\
= \lambda_1 H_{Q^1} & (X_1|X_0X_2) + \lambda_2 H_{Q^2}(X_1|X_0X_2)
\end{aligned}$$

where the inequality comes from the log sum inequality [25], with:

$$a_i = \lambda_i Q^i(x_0, x_1, x_2)$$

and

$$b_i = \lambda_i Q^i_{X_2}(x_2)$$

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.

2.4 Distributed power allocation case study

2.4.1 Case study description

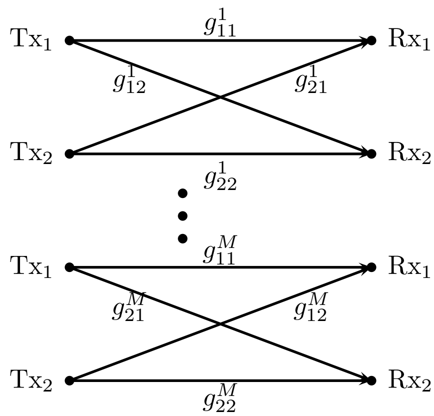


Figure 2.1: Case study considered in Sec. 2.4: an interference channel with 2 transmitters (Txs), 2 receivers (Rxs), and $M \geq 1$ non-overlapping frequency bands. One feature of the retained model is that both power allocation policies and channel gains g_{ij}^m are assumed to lie in finite discrete sets.

We now consider the specific problem of power allocation over $M-$band interference channels with two transmitter-receiver pairs. Transmission 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:

$$\mathcal{P}_i = \left\{ \frac{P_{\max}}{\ell} e_\ell : \ell \in \{1,\dots,M\}, e_\ell \in \{0,1\}^M, \sum_{i=1}^{M} e_\ell(i) = \ell \right\} \tag{2.22}$$

where $P_{\max}$ is the the power budget available at a transmitter. Each channel channel is assumed to lie 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. 2.3, the results derived therein can be applied here. At last, for a given time-slot, the instantaneous or the stage payoff function which is common to the transmitters is chosen to be:

$$w : \left| \begin{array}{ccc} \mathcal{G} \times \mathcal{P}_1 \times \mathcal{P}_2 & \to & \mathbb{R}^+ \\ (g, p_1, p_2) & \mapsto & \displaystyle\sum_{i=1}^{2} \sum_{m=1}^{M} B_m \log_2 \left(1 + \frac{g_{ii}^m p_i^m}{\sigma^2 + g_{-ii}^m p_{-i}^m} \right) \end{array} \right. \tag{2.23}$$

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.

2.4.2 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 [27] 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:

$$g_{ii}^1 \in \{0.1, 1.9\}, \quad i \in \{1,2\}$$
$$g_{ij}^2 \in \{0.15, 1.85\}, \quad (i,j) \in \{1,2\}.$$

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 the high interference regime corresponds to the $P((g_{ij}^2|i \neq j) = 1.85) = 1 - 0.1 = 0.9$, thus creating high interference due to higher probability for a greater value of $(g_{ij}^2|i \neq j)$ which is precisely the interference in the band 2. The similar intuition holds for low interference regime as well. 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 2.12;
- 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. 2.2 represents the gain allowed by asymmetric coordination w.r.t. the case where the transmitters always use the uniform power allocation policy (Blind Policy). This gain can be as high as 40% for the considered range of SNR. It is seen that the gain are particularly significant when the interference is high (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

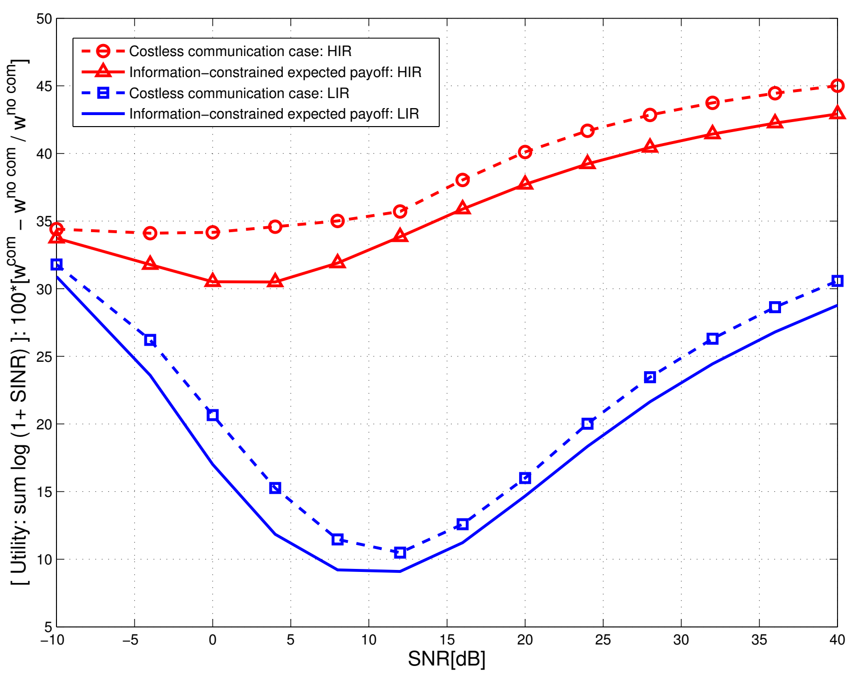


Figure 2.2: Relative gain in terms of expected payoff ("OP/BP - 1" in [%]) vs SNR[dB] obtained with the Optimal policy (OP) (with and without communication cost) when the reference policy is to put half of the power on each band (BP). Red curves correspond to the HIR, and blue curve to the LIR. $B_1 = B_2 = 10\text{MHz}$.

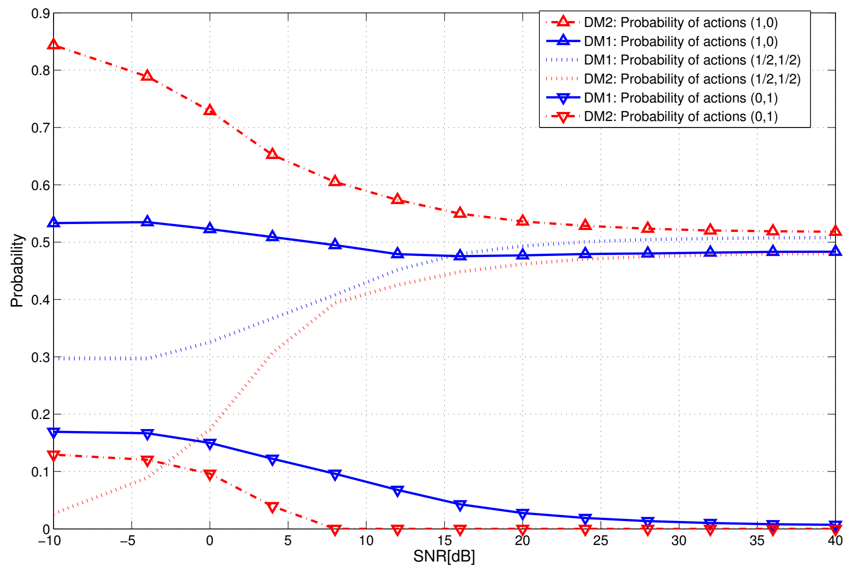


Figure 2.3: Marginal probability distributions $q_{X_1}(\cdot)$ $q_{X_2}(\cdot)$ of Transmitter 1 and Transmitter 2 for the optimal policy vs SNR[dB] for the High Interference Regime. $B_1 = B_2 = 10$MHz.

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. 2.3, 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. 2.3 that the probability of the action $(0, 1)$ (using the shared band) is zero from lower SNR values for DM2 than for DM1. This can be explained by the fact that the higher the power available for both DM, the higher the interference in the non-protected band. However, DM1 still chooses to play this action as it

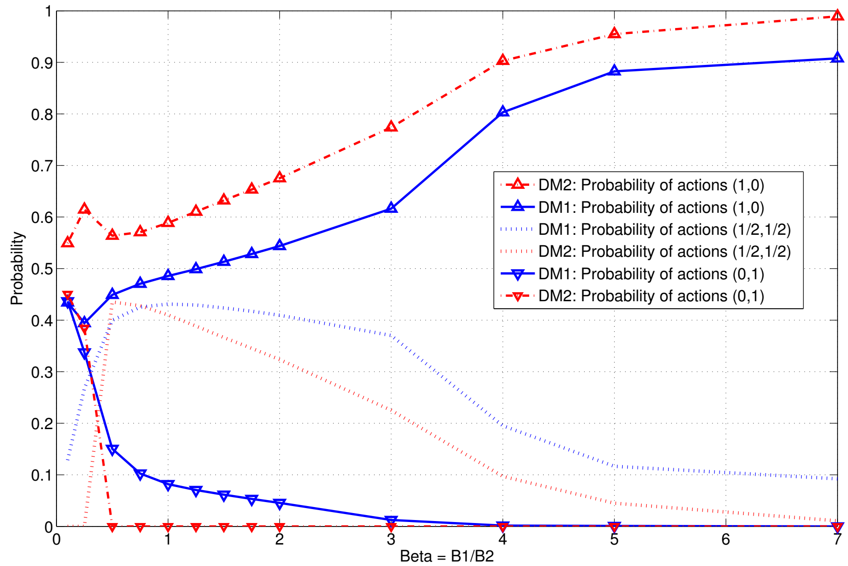


Figure 2.4: Marginal probability distributions $q_{X_1}(\cdot)$ $q_{X_2}(\cdot)$ of Transmitter 1 and Transmitter 2 for the optimal policy vs β for the optimal policy for the High Interference Regime, where $\beta = \frac{B_1}{B_2}$. For this simulation, we chose SNR=10[dB].

has knowledge of channel gains and can use the interference band to improve his utility. The same argument stands for Fig. 2.4.

Lastly, Fig. 2.4 shows the influence of the bandwidths on the power allocation policies. Not surprisingly, the higher the bandwidth of the protected band is, 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.

2.5 Conclusions

The analysis 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.

Chapter 3

Continuous Alphabets - Witsenhausen Cost Function

All means (even continuous) sanctify the discrete end.

Doron Zeilberger

In this chapter, we consider a team with two agents who are trying to maximize their common payoff over a long time period, i.e., composed of many time-slots. At every time-slot or stage $t \in \{1, \cdots, T\}$, Agent i, $i \in \{1, 2\}$ chooses its action $x_k \in \mathcal{X}_k$, where $\mathcal{X}_k$ is a *continuous* set. The instantaneous payoff function $w(x_0, x_1, x_2)$ depends on the realization of the random variable X_0 with realizations $x_0 \in \mathcal{X}_0$. The set $\mathcal{X}_0$ is also continuous and realizations of X_0 are assumed to be i.i.d.. The information structure is thus similar to the one considered in 2, albeit with alphabets for nature state $\mathcal{X}_0$ and agents' actions $\mathcal{X}_i$ being continuous.

A problem with the same information structure was addressed for the first time in [12]. Therein, the assumptions made are as follows: it is assumed that at any time Agent 1 knows the past, current, and future realizations of X_0 perfectly, whereas Agent 2 only observes the actions of Agent 1 in a *strictly* causal manner. Such a scenario has been extended in a couple of papers cited further.However, all of them treated the case where the action sets as well as the system state set were discrete and finite, i.e., $\forall\ k \in \{0, 1, 2\}, |\mathcal{X}_k| < \infty$. Reference [12] treated the case of an information structure in which Agent 2 has perfect observation and showed that the average performance characterization is equivalent to finding the appropriate information constraint. In [20] this result was generalized to the case where Agent 2 has imperfect observations. While all these contributions assume a strictly causal knowledge of the system state x_0 at Agent 2, the case where this assumption is relaxed was first presented in

[28] and treated rigorously in [21].

The main contribution of the work presented in this chapter is to generalize this approach of finding the limiting performance to the case of continuous action sets, which is an important case for control problems as many designs involve continuous controlers. A second contribution is to consider the Witsenhausen cost function [19] as a common cost function to be minimized by the $2-$agent team under the two mentioned scenarios in terms of information structures; this establishes for the first time a link between [12][20][21] and [19].

Indeed, although the Witsenhausen problem can be seen as a one-shot coordination problem, whereas we consider a long-term coordination problem here, the idea of joint control-communication strategies is present in both formulations. As it will be seen, characterizing the feasible performance of the long-term coordination problems amounts to determining a certain information constraint. Although information constraints normally appear when (large) sequences intervene, it has to been noted that one-shot problems closely related to the Witsenhausen problem have been solved by introducing an information constraint; this is the case, for instance, for the Gaussian Test Channel (GTC) [22. The approach we adopt has connections with that of [29] where probability distributions which minimize the cost function are used. However, in the latter the authors restrict their attention to what modifications render the (one-shot) Witsenhausen problem simpler to solve, and do not tackle the general framework of long-term implicit communication.

The chapter is structured as follows. Section 3.1 provides the proposed problem formulation. It explains that characterizing the feasible set of expected common payoffs amounts to characterizing implementable joint probability distributions. Section 3.2 provides, for the two information structures considered, the two information constraints which allows one to characterize the implementable distributions. The Gaussian case is provided as a special instance, which establishes a connection with the dirty-paper coding problem [30]. We also provide the information constraints for other continuous distribution in the case of where the second agent has perfect observation of first agent's actions. In Section 5.3.2, we discuss the Witsenhausen cost function in context to our problem. Section 3.3.1 describes the numerical analysis for the special instance of payoff function (which equals minus the Witsenhausen cost function) and provide numerical results. Section 3.4 concludes the chapter.

3.1 Problem statement

Consider two agents, Agent 1 and Agent 2, who want to coordinate through their actions $x_1 \in \mathcal{X}_1$ and $x_2 \in \mathcal{X}_2$. The problem is said to be distributed in the sense that each agent can only control one variable of their common payoff function $w(x_0, x_1, x_2)$. The action set for both agents $\mathcal{X}_1$, $\mathcal{X}_2$ as well as the set of system states $\mathcal{X}_0$ are continuous sets. The realizations of the system state are assumed to be i.i.d. and generated from a random variable X_0 whose probability density function is denoted by $f_0(x_0)$. We shall use the notation $f_V(v)$ or $f(v)$

to refer to the probability density function of the generic continuous random variable V. The control strategies of Agents 1 and 2 are sequences of functions which are respectively defined by:

$$\left\{ \begin{array}{lccc} \sigma_t & : & \mathcal{X}_0^T & \rightarrow & \mathcal{X}_1 \\ \tau_t^a & : & \mathcal{X}_0^{t-1} \times \mathcal{Y}^{t-1} & \rightarrow & \mathcal{X}_2 \\ \tau_t^b & : & \mathcal{Y}^{t-1} & \rightarrow & \mathcal{X}_2 \end{array} \right. \tag{3.1}$$

where $T \geq 1$ is the total number of stages over which the agents are assumed to interact, $\mathcal{Y}$ is the observation set of Agent 2 and the superscripts a or b correspond to the two considered scenarios in terms of observation structure.

The control strategy for Agent 1 σ_t basically means that it knows the realizations of the system state for all T beforehand, and uses that information to choose its actions; note that the methodology used in this chapter can also be exploited under less restrictive knowledge assumptions at Agent 1. The merit of the assumptions made for Agent 1 is that it allows one to make progress in the direction of quantifying the relationship between agents' observation capabilities and reachable performance, which is not well understood. Additionally, there already exist applications for which it is relevant: coordination between robots when a leader knows the trajectory in advance; distributed power control in wireless networks; robust image watermarking. For Agent 2, we consider two different control strategies τ_t^a and τ_t^b. The control strategy of *scenario a* assumes that Agent 2 observes the past realizations of the system state $x_0(1), ..., x_0(i-1)$ as well as $y(1), ..., y(i-1)$. The control strategy of *scenario b* is only based on the latter sequence and seems to be more in line with a possible information structure of a long-term version of the Witsenhausen problem. In any case, it is less demanding in terms of information assumptions. The observations $y(1), ..., y(T)$ are assumed to be generated by a memoryless channel whose transition probability is denoted by γ and verifies a Markov condition $f_{Y|X_0,X_1,X_2}(y|x_0, x_1, x_2) = \gamma(y|x_1)$. The additive white Gaussian noise (AWGN) channel $Y = X_1 + Z$ is an intensively used model which verifies this condition.

The instantaneous team payoff function is denoted by $w(x_0, x_1, x_2)$. Since X_0 is not deterministic we shall be considering the expected payoff

$$\mathbb{E}_f[w(X)] = \int_{x \in \mathcal{X}} w(x_0, x_1, x_2) f(x_0, x_1, x_2) \mathrm{d}x_0 \mathrm{d}x_1 \mathrm{d}x_2 \tag{3.2}$$

where $X = (X_0, X_1, X_2)$, $x = (x_0, x_1, x_2)$, and $\mathcal{X} = \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2$. In the sequel, we will also denote by $W(f)$ the above expected payoff i.e., $W(f) = \mathbb{E}_f[w(X)]$.

What matters for the expected payoff is function f which characterizes the possible correlations among the three random variables X_0, X_1, and X_2. This correlation precisely measures the degree to which the agents can coordinate with each other and the system state. To understand the relationship between the agents strategies (5.8) and the expected payoff (3.2), let us define the notion of implementable distributions.

Definition 2 (Implementability)**.** *Let $s \in \{a, b\}$ be the assumed information structure. The probability density function $\overline{f}(x_0, x_1, x_2)$ is implementable if there exists a pair of control strategies $(\sigma_t, \tau_t^{\mathrm{s}})$ such that as $T \to +\infty$, we have for all $(x_0, x_1, x_2) \in \mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2$,*

$$\frac{1}{T}\sum_{i=1}^{T}\int_{y\in\mathcal{Y}} f_{X_0X_1X_2Y,i}(x_0, x_1, x_2, y) \to \overline{f}(x_0, x_1, x_2) \tag{3.3}$$

where $f_{X_0X_1X_2Y,i} = \gamma \times f_{X_1,X_2|X_0,i} \times f_0$ is the joint distribution induced by (σ_t, τ_t^s) at stage i.

Note that since the expectation value of the payoff is a linear operator with respect to the distribution f, the time averaged expected payoff $\overline{W}$ is reachable if and only if the corresponding distribution $\overline{f}$ is implementable. In Section 3.2, we shall characterize the set of reachable or feasible average payoffs under the information structure given by (5.8), which is equivalent to characterizing the set of implementable distributions.

3.2 Performance analysis: limiting performance characterization

3.2.1 General case

In the case of finite alphabets $|\mathcal{X}_i| < \infty$, $i \in \{0, 1, 2\}$, it has been shown in the cases which have been treated so far [12], [20], [28], [21] that characterizing the set of implementable (mass) probability distributions amounts to determining a certain information constraint. For instance in the previous chapter, i.e. case of *discrete* sets and *perfect observation* ($Y = X_1$), the necessary and sufficient condition for a joint probability mass distribution $Q(x_0, x_1, x_2)$ to be implementable is that

$$H_Q(X_0) + H_Q(X_2) - H_Q(X_0, X_1, X_2) \leq 0 \tag{3.4}$$

where H_Q is the discrete entropy function under a fixed joint distribution (see Section [31] for the different expressions of the entropy used in this section).

A well-known reasoning in information theory [31], and intensively used in control when communication problems are involved, is to use the information constraint derived in the discrete case and just replace the discrete entropy function with the differential entropy. It can be proved that this reasoning is perfectly valid if considered continuous variables are Gaussian (see e.g., [32] for a recent reference). For coordination problems such as the one under investigation, imposing the agents' actions to be Gaussian is generally suboptimal. Elaborating further, if we replace the discrete entropy function with the differential entropy we obtain

$$h_f(X_0) + h_f(X_2) - h_f(X_0, X_1, X_2) \leq 0 \tag{3.5}$$

where h_f is the differential entropy under the fixed joint distribution f.

It turns out that this condition can be shown to be non-necessary in general, indicating that the transition from the discrete case to the continuous case needs some special care in the problem under investigation. To convince the reader, let us recall one of the Cantor's theorems (see e.g., [33]). There exists a bijective map from $\mathbb{R}^T$ to $\mathbb{R}$. Therefore, a possible control strategy for Agent 1 might be as follows. On the first stage, Agent 1 maps or encodes the whole sequence of states $(x_0(1), \cdots, x_0(T)) \in \mathbb{R}^T$ into a single action $x_1(1) \in \mathbb{R}$. Since Agent 2 observes this action perfectly, it can decode it perfectly and is thus informed of the sequence of states as well. This would mean that from stage $i = 2$, the two agents can correlate their actions in a arbitrary manner with the system state; in particular they can choose the pair (or one of the pairs) which maximizes w at a given stage $i \geq 2$. This means that any probability density function $\overline{f}_{X_0 X_1 X_2}$ can be implemented (asymptotically), contradicting the fact that any implementable distribution has necessarily to verify the continuous counterpart of (3.4) which is (3.5). This apparent contradiction comes from the fact that expressing (3.4) in the continuous case with differential entropies relies on assumptions which need to be specified rigorously for the problem. Indeed, the information constraint can be shown to be necessary and sufficient for implementability within some classes of random variables. One of the broadest classes which is known is provided in [34]. It turns out that if one wants to define a probability measure on the Cantor set, one does not fall into this broad class which is specified below. Let's first give the definition of a field in probability theory.

Definition 3 (field)**.** *Let $(\Omega, \mathcal{B})$ be a measurable space. We call field $\mathcal{F}$ a collection of subset of Ω such that :*

$$\Omega \in \mathcal{F}$$

$$if\ F \in \mathcal{F}\ then\ F^c \in \mathcal{F}$$

$$\mathcal{F}\ is\ stable\ under\ finite\ union$$

A set A of a field $\mathcal{F}$ is called an atom if and only if the only subsets which are also member of the field are the set itself and the empty set.

Definition 4 (base)**.** *A sequence of finite field $\mathcal{F}_n$; $n = 0, 1, ...$ is called a basis of a field $\mathcal{F}$ if $\mathcal{F}_n \uparrow \mathcal{F}$ and if G_n is a sequence of atoms of $\mathcal{F}_n$ such that $G_n \in \mathcal{F}_n$ and $G_{n+1} \subset G_n$, $n = 0, 1, 2, ...$ then $\cap_{n=1}^{\infty} G_n \neq \emptyset$.*

A sequence $\mathcal{F}_n$; $n = 0, 1, ...$ is called a basis of a measurable space $(\Omega, \mathcal{B})$ if $\mathcal{F}_n$ are a basis of a field $\mathcal{F}$ which generates $\mathcal{B}$: $\mathcal{B} = \sigma(\mathcal{F})$. A field $\mathcal{F}$ is called standard if it has a basis. A measurable space $(\Omega, \mathcal{B})$ is called standard if it can be generated by a standard field i.e. $\mathcal{B}$ has a basis. We can now define the mutual information in a standard space provided by [35].

Definition 5 (Mutual Information)**.** *Let $(\Omega, \mathcal{F}, P)$ be a standard probability space and $X \in \mathbb{R}$, $Y \in \mathbb{R}$ two generic random variables: $X : \Omega \to \mathcal{A}_X$, $Y : \Omega \to$*

$\mathcal{A}_Y$ with $(\mathcal{A}_X, \mathcal{B}_X)$, $(\mathcal{A}_Y, \mathcal{B}_Y)$ two measurable spaces. Let $\mathcal{F}_X = X^{-1}(\mathcal{A}_X)$ and $\mathcal{F}_Y = Y^{-1}(\mathcal{A}_Y)$ the sub-σ-algebra of $\mathcal{F}$ induced by X and Y. Let

$$\mathcal{P}_X = \{A_j\}_{j=1}^{N_X} \subset \mathcal{F}_X \text{ and } \mathcal{P}_Y = \{B_j\}_{j=1}^{N_Y} \subset \mathcal{F}_Y \tag{3.6}$$

be finite partitions of Ω. With these partitions we associate the following random variables:

$$\begin{array}{lll} \widetilde{X}(\omega) = j & \text{for } \omega \in A_j & \text{with } 1 \leq j \leq N_X \\ \widetilde{Y}(\omega) = j & \text{for } \omega \in B_j & \text{with } 1 \leq j \leq N_Y \end{array}. \tag{3.7}$$

The mutual information between $X \in \mathbb{R}$ and $Y \in \mathbb{R}$ is then defined by:

$$i(X;Y) = \sup_{\mathcal{P}_X, \mathcal{P}_Y} I(\widetilde{X};\widetilde{Y}) \tag{3.8}$$

where I is the classical mutual information between two discrete random variables [25]. Similarly, the conditional mutual information is defined by

$$i(X;Y|Z) = i(X;Y,Z) - i(X;Z). \tag{3.9}$$

The above framework is exploited to prove the following two theorems.

Theorem 2 (Scenario a)**.** *Assume that all random variables under use are defined on a standard probability space. Consider a joint probability density distribution $\overline{f}(x_0, x_1, x_2)$ such that $\forall x_0 \in \mathcal{X}_0, \int_{x_1,x_2} \overline{f}(x_0, x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 = f_0(x_0)$. Then, the distribution $\overline{f}$ is implementable if and only if $f(x_0, x_1, x_2, y)$ verifies the following information constraint:*

$$i_f(X_0;X_2) \leq i_f(X_1;Y|X_0,X_2) \tag{3.10}$$

where the arguments of the mutual information $i_f(.)$ are defined from f and $f(x_0, x_1, x_2, y) = \overline{f}(x_0, x_1, x_2)\gamma(y|x_1)$.

Theorem 3 (Scenario b)**.** *Assume that all random variables under use are defined on a standard probability space. Consider a joint probability density distribution $\overline{f}(x_0, x_1, x_2)$ such that $\forall x_0 \in \mathcal{X}_0, \int_{x_1,x_2} \overline{f}(x_0, x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 = f_0(x_0)$. Then, the distribution $\overline{f}$ is implementable if and only if $f(x_0, x_1, x_2, y, x_1')$ verifies the following information constraint:*

$$i_f(X_0;X_2) \leq i_f(X_1';Y,X_2) - i_f(X_1';X_0,X_2) \tag{3.11}$$

with $f(x_0,x_1,x_2,y,x_1') = f_{X_1'|X_0X_1X_2}(x_1'|x_0,x_1,x_2)\gamma(y|x_1)\overline{f}(x_0,x_1,x_2)$. X_1' is an auxilary variable which helps us exploit the joint typicality between X_1' and X_1 as well as X_1' and X_2 to create coding and decoding schemes.

We see that the first theorem provides a full characterization of implementable densities in scenario a. The second theorem, which relies in part on Gel'fand and Pinsker coding [36], provides a sufficient condition for implementability in scenario b; note that, as originally done in [36], we introduce an auxiliary random variable X_1' to describe the information constraint. These theorems therefore allow one to know to what extent a team can coordinate under the assumed information structure. In general, to determine the ultimate performance in terms of average payoff, an optimization problem for the functional $W(\overline{f})$ has to be solved. The constraints are that: $\overline{f}$ has to be a density function; its marginal over (x_1, x_2) has to be f_0; the density f as defined through the considered theorem has to meet the information constraint. In the next subsection, we apply the derived general result to a special case of probability distributions namely, Gaussian probability density functions. This allows one to exhibit a case where the information constraints can be quite easily expressed and to establish an interesting link with the work by Costa on dirty-paper coding [30].

3.2.2 Gaussian case

Here, we assume that all variables which intervene in the information constraints are Gaussian. Agent 2 is assumed to observe the actions of Agent 1 through an additive white Gaussian noise channel: $Y = X_1 + Z$ with $Z \sim \mathcal{N}(0,1)$. Let σ_0^2, σ_1^2, and σ_2^2 respectively denote the variances of X_0, X_1, and X_2. The correlation coefficient between X_i and X_j, $i \neq j$ is denoted by ρ_{ij}. Using these notations and specializing (3.10) and (3.11) in the Gaussian case the following results can be proved; proofs are omitted here for obvious space limitations.

Proposition 4 (Information constraint in scenario a)**.** *Fix σ_0^2. A necessary and sufficient condition for a joint probability density function $\overline{f}_{X_0X_1X_2}$ to be implementable is that the variances and correlation coefficients verify the following inequality:* $-(\sigma_2^2\rho_{01}^2 - 2\rho_{01}\rho_{02}\rho_{12} + \sigma_1^2\rho_{02}^2 + \sigma_0^2\rho_{12}^2 - \sigma_0^2\sigma_1^2\sigma_2^2) \times (\sigma_2^2\rho_{01}^2 - 2\rho_{01}\rho_{02}\rho_{12} + \sigma_1^2\rho_{02}^2 + \sigma_0^2\rho_{12}^2 - \sigma_0^2\sigma_1^2\sigma_2^2 + \rho_{02}^2) \leq 0.$

Proposition 5 (Information constraint in scenario b)**.** *Fix σ_0^2. Let $\alpha_2 \in \mathbb{R}$. A sufficient condition for a joint probability density function $\overline{f}_{X_0X_1X_2}$ to be implementable is that the variances and correlation coefficients verify the following inequality:* $(-\rho_{01}^2 + \sigma_0^2 + \sigma_0^2\sigma_1^2) \times (\sigma_2^2\rho_{01}^2 - 2\rho_{01} + 2\alpha_2\rho_{12} + \sigma_1^2\rho_{02}^2 + \sigma_0^2\rho_{12}^2 - \sigma_0^2\sigma_1^2\sigma_2^2) - \sigma_0^2\sigma_2^2(\rho_{01}^2 - \sigma_0^2\sigma_1^2 + \alpha_2^2\rho_{02}^2 - \alpha_2^2\sigma_0^2\sigma_2^2 + \alpha_2^2\sigma_1^2\rho_{02}^2 + \alpha_2^2\sigma_2^2\rho_{01}^2 + \alpha_2^2\sigma_0^2\rho_{12}^2 - 2\alpha_2\sigma_0^2\rho_{12} + 2\alpha_2\rho_{01}\rho_{02} - \alpha_2^2\sigma_0^2\sigma_1^2\sigma_2^2 - 2\alpha_2^2\rho_{01}\rho_{02}\rho_{12}) \leq 0.$

For finding the information constraint for scenario b, we have assumed that $X_1^{'} = X_1 + \alpha_0 X_0 + \alpha_2 X_2$. This choice is inspired by the Costa's dirty paper coding scheme [30]. The parameter α_2 is related to the choice we made for the auxiliary variable in Theorem 3 .

As seen through Proposition 5, the value of the parameter α_0 does not play any role in the constraint, showing that only the correlation level between the agents' actions X_1 and X_2 has to be tuned properly. The inequality constraint function of Proposition 5 can be shown to be strictly convex w.r.t. α_2 and the optimum point $\alpha_2^\star$ is given by:

$$\alpha_2^\star = \frac{\sigma_0^2\rho_{12} - \rho_{01}\rho_{02}}{\rho_{02}^2 - \sigma_0^2(\sigma_2^2 - \rho_{12}^2 + \sigma_1^2\sigma_2^2) + \sigma_1^2\rho_{02}^2 + \sigma_2^2\rho_{01}^2 - 2\rho_{01}\rho_{02}\rho_{12}}. \quad (3.12)$$

When the communication signal-to-noise ratio SNR $= \frac{\mathbb{E}(X_1^2)}{\mathbb{E}(Z^2)} = \sigma_1^2 \to \infty$, it is seen that $\alpha_2^\star \to 0$ and the choice $X_1^{'} = X_1$ is optimal. When SNR $\to 0$, we see that $\alpha_2^\star \to \frac{\sigma_0^2\rho_{12} - \rho_{01}\rho_{02}}{(\rho_{02}^2 - \sigma_0^2\sigma_2^2) + \sigma_0^2\rho_{12}^2 + \sigma_2^2\rho_{01}^2 - 2\rho_{01}\rho_{02}\rho_{12}}$.

As an aside, using the Cauchy-Shwarz Inequality one can show that for any 3 correlated random variables, the bounds for ρ_{12} are $\frac{\rho_{01}\rho_{02}}{\sigma_0^2} - \frac{\sqrt{(\sigma_0^2\sigma_2^2 - \rho_{02}^2)(\sigma_0^2\sigma_1^2 - \rho_{01}^2)}}{\sigma_0^2} \leq \rho_{12} \leq \frac{\rho_{01}\rho_{02}}{\sigma_0^2} + \frac{\sqrt{(\sigma_0^2\sigma_2^2 - \rho_{02}^2)(\sigma_0^2\sigma_1^2 - \rho_{01}^2)}}{\sigma_0^2}$. Thus, one sees that when $\rho_{01} = \sigma_0\sigma_1$ and $\rho_{02} = \sigma_0\sigma_2$, we obtain $\rho_{12} = \frac{\rho_{01}\rho_{02}}{\sigma_0^2}$. This is true for all SNR. Incidentally this is also the condition for both the numerator and the denominator of α_2^* in (3.12) to be zero.

In the case of perfect monitoring, i.e. $Y = X_1$, we can show that the information constraint suitably simplified is in fact convex atleast for the special case of $X_0 = X_1$. (Proof in 3.5.1).

3.2.3 Other Distributions

Since Theorems 2, 3 can be applied to the random variables following any distribution, it is worthwhile to explicit the information constraints for distributions other than normal. However, both scenarios involve imperfect observations by the second agent. Since the noise model is AWGN, it is difficult to find closed form expressions of the information constraint for the variables not jointly distributed normally. Therefore, in this section, we consider the noiseless case, i.e. $Y = X_1$. We consider some other standard distributions which might be useful in modelling other co-ordination problems.

Another difficulty arises due to the inexistence of closed form expressions for entropy for most joint continuous distributions. Amongst the few distributions which do, we also need the property that marginalization w.r.t. any variable gives back the same joint distribution, failing which we might again have the problem of no closed form expression for entropy. Two common distributions which satisfy these constraints (Sec. 3.5) are Multivariate Burr and Multivariate Exponential.

Proposition 6 (Multivariate Burr)**.** *When the distribution $f(x_0, x_1, x_2)$ is a Multivariate Burr with the following pdf (Sec. 3.5):*

$$f(x_0, x_1, x_2) = \frac{\Gamma(\alpha+3)}{\Gamma(\alpha)} c_0 d_0 c_1 d_1 c_2 d_2 \frac{x_0^{c_0-1} x_1^{c_1-1} x_2^{c_2-1}}{(1 + d_0 x_0^{c_0} + d_1 x_1^{c_1} + d_2 x_2^{c_2})^{\alpha+3}} \quad (3.13)$$

where $\alpha, c_i, d_i, x_i > 0$*,* $i \in \{0, 1, 2\}$

then the Information constraint in the case of perfect monitoring can be written as:

$$log\frac{(\alpha+1)(\alpha+2)}{\alpha} - \frac{\alpha^3+7\alpha^2+10\alpha+2}{\alpha(\alpha+1)(\alpha+2)} + log c_1 d_1^{1/c_1} - \frac{c_1-1}{c_1}\{\psi(\alpha)-\psi(1)\} \leq 0 \tag{3.14}$$

where $\psi(\cdot)$ *is the standard psi function as defined in mathematics literature.*

It is interesting to note that the distribution parameters c_0, d_0, c_2, d_2 do not intervene in the information constraint.

Proposition 7 (Multivariate Exponential)**.** *When the distribution* $f(x_0, x_1, x_2)$ *is a Multivariate Burr with the following pdf (Sec. 3.5):*

$$f(x_0, x_1, x_2) = \frac{\frac{\alpha}{\theta_0} e^{\frac{x_0-\lambda_0}{\theta_0}} \frac{\alpha+1}{\theta_1} e^{\frac{x_1-\lambda_1}{\theta_1}} \frac{\alpha+2}{\theta_2} e^{\frac{x_2-\lambda_2}{\theta_2}}}{(e^{\frac{x_0-\lambda_0}{\theta_0}} + e^{\frac{x_1-\lambda_1}{\theta_1}} + e^{\frac{x_2-\lambda_2}{\theta_2}} - 1)^{\alpha+3}} \tag{3.15}$$

where $x_i > \lambda_i$*,* $\alpha, \theta_i > 0$*,* $i \in \{0, 1, 2\}$

then the information constraint in the case of perfect monitoring can be written as:

$$log\frac{(\alpha+1)(\alpha+2)}{\alpha\theta_1} - \frac{(\alpha^2+7\alpha+7)}{(\alpha+1)(\alpha+2)} \leq 0 \tag{3.16}$$

Again, interestingly, like in the case of multivariate Burr, the parameters θ_0 and θ_2 disappear from the information constraint.

3.3 Application to the Witsenhausen cost function

An important 2-Agent co-ordination problem with continuous action alphabets is the Witsenhausen counterexample (See **??** for a detailed description). The information structure of the problem is similar to the one considered by us. One agent has perfect observation of the nature state, whereas the second agent has noisy observations of the actions of the first agent. For this reason, we consider the Witsenhausen cost (times minus one) as the payoff function to be minimized over a large number of stages and reducing it to a convex optimization problem. We find the limiting performance in terms of coordination for the two-agent team.

However, note that this is significantly different from the original Witsenhausen's counterexample [19] as we optimize over many stages with one agent 'knowing' the future realizations of the nature state X_0. In the original counterexample, the 'informed' agent only knows the current realization of X_0. Thus, the problem has only one stage and the aim is to minimze the expected cost for that stage. The problem has been of great interest since it was proposed in 1968, especially because it was one of the first examples where actions played a dual role of communicating as well as optimizing, rendering the solution of the optimization problem non trivial. The specific payoff function function we consider in this section equals minus the Witsenhausen cost function that is,

$$w(x_0, x_1, x_2) = -\left[k(x_0 - x_1)^2 + (x_1 - x_2)^2\right], \ k \geq 0. \tag{3.17}$$

Note that the notations used here differ from those used in [19].

Although this cost function is inspired by Witsenhausen's Counterexample, there are a few very important differences between the application of our theory to the Witsenhausen's Cost function and the Witsenhausen's counterexample. First, we optimise the average cost over a large time period, unlike in Witsenhausen's Counterexample where the one shot expectation value for the cost is minimised. Second, we assume strictly causal knowledge at Agent 2, whereas this is not the case in the original counterexample. To apply our approach to the original problem, albeit in the repeated case, one will need to consider a new scenario with the causality condition relaxed. We tackle the orginial counterexample in the second part of this thesis after developing the theory for the information structure corresponding to Witsenhausen Counterexample.

3.3.1 Numerical analysis

While the optimization problem is well defined, finding the joint distribution $f_{X_0X_1X_2Y} = \gamma(y|x_1) * f_{X_1X_2|X_0} * f_0$ is not a trivial task. Eventhough f_0 is defined by the problem and $\gamma(y|x_1)$ can be generated given the noise model, we still need to find $f_{X_1X_2|X_0}$, and since the search space is over all possible distributions, the computational complexity of such a search is very high. Therefore, in this section we restrict ourselves to two cases which are simpler to handle, complexity wise, but might be sub-optimal.

Both the cases use the results proven in previous sections for the continuous variables. However for the first case, we use a quantizer to discretise $(\mathcal{X}_0, \mathcal{X}_1, \mathcal{X}_2, \mathcal{Y})$ and optimise over them. This approach is inspired by the success of quantizers in finding better solutions for the original Witsenhausen Counterexample. To compare with a strategy most resembling a 'linear control' strategy, we choose the case where all variables are supposed to be gaussian, for which we had simplified the information constraint in terms of variances and covariances in section 3.2.2. Clearly, it would be interesting to compare the two strategies, to see whether like for Witsenhausen Counterexample, non linear (discretisation) strategies outperform 'linear' (and thus continuous) strategies. This argument is used to motivate our choices of simulations, but the similarities are not so straightforward.

The following parameters are taken to be given for the problem: $k = 1, \sigma_0^2 = 25$, and $\mathbb{E}z^2 = 1$ and are common for both simulations (unless specified otherwise). Also, the information constraint considered is for scenario a.

Discrete case : we quantize all random variables to take nine values: $\mathcal{X}_0 = \mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y} = \{-\frac{24}{7}\sigma_0, -\frac{16}{7}\sigma_0, -\frac{8}{7}\sigma_0, -\frac{3}{7}\sigma_0, 0, \frac{3}{7}\sigma_0, \frac{8}{7}\sigma_0, \frac{16}{7}\sigma_0, \frac{24}{7}\sigma_0\}$, with $\sigma_0 = 5$. Indeed, as we consider the continuous random variable $\tilde{X}_0 \sim \mathcal{N}(0, \sigma^2)$, we partition uniformly the continuous space that $\tilde{X}_0$ is defined over, so that 99.99 % of the probability mass function lies in the chosen interval. It is well known that considering the interval $[-4\sigma, 4\sigma]$ achieves this. We then calculate the transition probabilities $P(Y|X_1)$ where $\tilde{Y} = X_1 + Z$, Z is supposed to be a Gaussian random variable: $Z \sim \mathcal{N}(0, 1)$, and Y is the discrete random variable that corresponds to $\tilde{Y}$. This problem is computationally simpler as it is easy to calculate entropies for discrete random variables. The optimization problem can be solved using convex optimization algorithms. We search for the joint distribution over $\mathcal{X}_0 \times \mathcal{X}_1 \times \mathcal{X}_2$ which minimizes the expectation of the Witsenhausen cost function. This approach is similar to the one used in [37], except for the cost function and the information constraint which takes into account the observation noise for Agent 2. It gives us an approximation to the continuous case and valuable ideas as we will explain now.

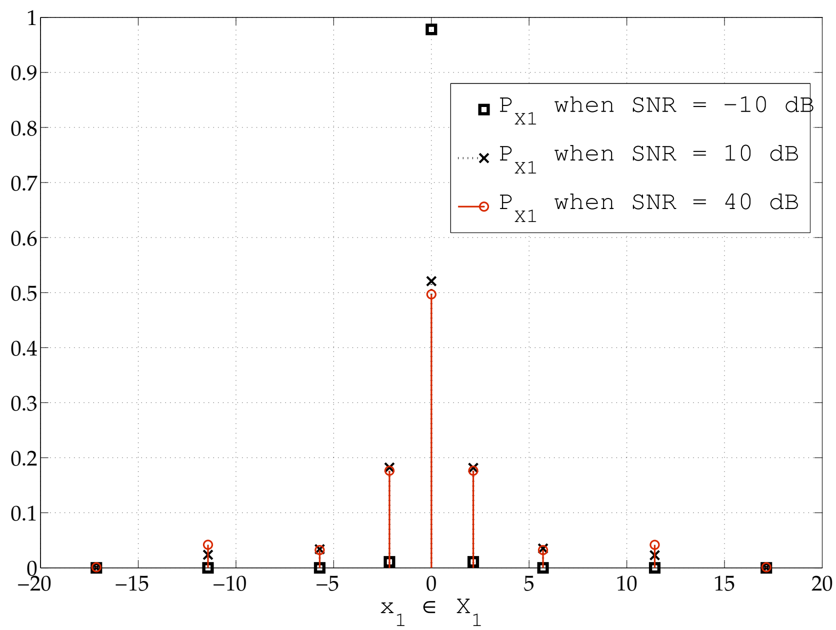


Figure 3.1: Optimal probability distribution of X_1 at different maximum SNRs. SNR=40 dB is the optimal distribution globally

Results : In Fig 3.1, for low SNR (-10dB), the probability is almost 1 for $X_1 = 0$. This is logical as $\sigma_1^2 = 0.1$ and thus Agent 1 does not have too much of a choice. For medium SNR (10 dB), we see the probabilities diverging slightly and resembling a Gaussian distribution. The same distribution is observed from SNR = 14 dB onwards as this leads to minimum cost. This can be seen from the graph of expected payoff vs SNR, Fig 3.2. At 40 dB, one sees a distribution with higher variance but whether it can be a gaussian is tough to speculate given the lack of points.

The salient feature of this approach is that it does not suppose any distribution for the variables a priori, and thus finds the optimal distribution, which is not necessarily Gaussian. However, it only searches for finite action alphabets, thus not attaining optimality in the general continuous case.

Gaussian Case : Guessing the optimal distributions to be Gaussian, we find a feasible set of variances and covariances which satisfy the information constraint calculated in Section 3. The feasible set is found by quantizing the search space for all the parameters. The search space are as follows: $\frac{\sigma_1^2}{\sigma_z^2} \in (-10, 40)_{dB}$, $\frac{\sigma_2^2}{\sigma_z^2} \in (-20, 13)_{dB}$, and $\rho_{01}, \rho_{12}, \rho_{02} \in (0, (\sigma_0\sigma_1, \sigma_1\sigma_2, \sigma_2\sigma_0))$. The other constraints of the optimization problem are trivially satisfied since we are taking all distributions to be probability distributions from the beginning.

For a given set of variance and covariance values which satisfy the information constraint, we find the expected payoff by evaluating the integral $\int_{x\in\{\mathcal{X}_0\times\mathcal{X}_1\times\mathcal{X}_2\}} f(x)w(x)\mathrm{d}x$. We do so using Monte Carlo simulations by randomly generating x_0, x_1 and x_2 (100000 draws) which follow the joint distribution $f(\mathbf{x})$ and averaging over the cost for the randomly generated triplets. We search over all the elements of the feasible set exhaustively to find the optimal joint Gaussian distribution which optimizes the Witsenhausen cost function.

Results : In Fig 3.2, we see that as $\mathbb{E}\frac{\sigma_1^2}{\sigma_z^2}$ (which could be looked at as Signal to Noise Ratio (SNR)) increases, the expected cost reduces initially and then becomes constant. This is because on the x-axis, we are considering the maximum SNR available, and in both the cases, one sees that after a certain SNR*, the payoff remains constant as the agents choose strategies with SNR*.

We notice that the discrete strategy does better than the continuous gaussian strategy and while this is not conclusive proof, it leads us to suspect that discretisation does better than continuous alphabets, which would be similar to the original Witsenhausen problem where discrete Non-Linear strategies were shown to outperform the best affine strategies. Although this might just be an artefact of gaussian variables being sub-optimal distributions for our problem.

For the Gaussian case, the optimal $(\sigma^*, \rho^*) = (\sigma_1^*, \sigma_2^*, \rho_{01}^*, \rho_{02}^*, \rho_{12}^*)$ was found to be $(4.6, 5.3, 4.8, 5.1, 4.9)$ and the optimal expected cost $\mathbb{E}(w)^*$ associated with it to be 18.64. The information constraint is numerically found to be saturated for the optimal point. Note that the optimal correlations ρ^* are not at their maximum values, given by $\sigma_i\sigma_j \geq \rho_{ij}$. This is precisely the information constraint which prevents the correlations to be at their maximum values, thus penalising future communication by Agent 1 and preventing the agents to coordinate better.

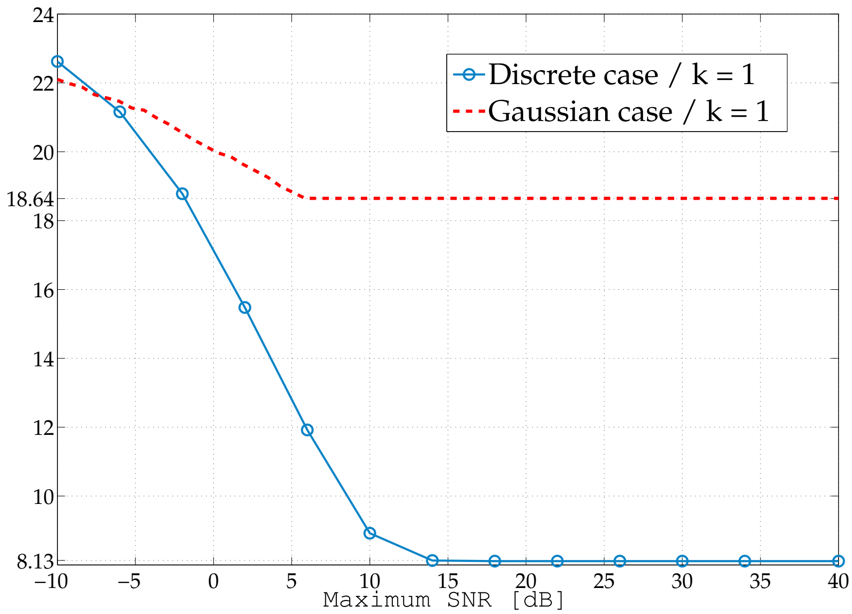


Figure 3.2: Scenario a; Plot of Expected payoff vs SNR for both the cases. One sees that the discrete strategy quickly beats the all-gaussian strategy, and achieves a much lower optimal cost

Ideally one should be able to do more extensive simulations for the discrete case, so as to get a better idea of the optimal distribution which could then be tackled using an approach similar to the gaussian case. However, the computational complexity prevents us from doing so currently.

3.4 Conclusions

We generalized the information constraint for the scenarios discussed in the chapter from a discrete case to a continuous case, showing equivalence between implementable distributions and reachable payoffs. This result is independent of the cost function, but it depends on the strategies and information structures. Thus we created a general framework for tackling problems with similar information structures, as well as showed a method of generalizing information constraint found for discrete cases to continuous ones. We also found an interesting link between scenario b and [30] which needs to be further explored. Since Witsenhausen cost function is an important area for research, we applied our framework to optimize this cost function in our scenario. While we could not provide simulations which solved the general optimization problem described in the chapter, we did gain some insights by simplifying the problem and reducing the computational complexity.

Further possibilities of exploration include better simulations with other types of distributions to see if they do better than Gaussian distributions, quantizing the alphabets with more points so as to approach continuity, as well as proving our results for other type of strategies and information structures. To compare a repeated version of Witsenhausen counterexample, one would need performance limit characterization with a different Information structure, similar to ones discussed in [9].

3.5 Appendix

3.5.1 Proof of Convexity of Information constraint for the Gaussian Case

The information constraint in the case of perfect monitoring can be given by substituting $Y = X_1$ in the theorem 2:

$$H_f(X_0) + H_f(X_2) - H_f(X_0, X_1, X_2) \leq 0 \tag{3.18}$$

where f is the joint distribution $f(x_0, x_1, x_2)$. Assume f to be a multivariate Normal distribution whose pdf is given by:

$$f(\mathbf{x}) = \frac{1}{(2\pi|\Sigma|)^{\frac{1}{2}}} e^{\frac{1}{2}\mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x}} \tag{3.19}$$

Σ is the covariance matrix $\Sigma_{ij} = \mathcal{E}X_iX_j$ Without loss of generality, we can assume the mean vector to be zero as the entropy does not depend on the mean. The entropy for the distribution f is:

$$H_f(X_0, X_1, X_2) = \frac{1}{2}\log(2\pi e|\Sigma|) \tag{3.20}$$

Using (3.18) and (3.20) and simplifying, we get the following conditions for the parameters of the distribution f:

$$P_0P_2 + 2\pi e(P_0\rho_{12}^2 + P_1\rho_{02}^2 + P_2\rho_{01}^2) - P_0P_1P_2 - 2\rho_{01}\rho_{02}\rho_{12} \leq 0 \tag{3.21}$$

where ρ_{ij} is the correlation between variables $X_i, X_j, i \neq j, \{i,j\} \in \{0,1,2\}$. and P_i is the variance of the variable X_i.

To check the convexity of the Information Constraint, we calculate the Hessian of the function w.r.t. the control variables $\{P_1, P_2, \rho_{01}, \rho_{02}, \rho_{12}\}$ in that order. The symmetric Hessian matrix can be expressed as follows:

$$\mathcal{H}(IC) = \begin{bmatrix} 0 & -2\pi e P_0 & 0 & 4\pi e\rho_{02} & 0 \\ -2\pi e P_0 & 0 & 4\pi e\rho_{01} & 0 & 0 \\ 0 & 4\pi e\rho_{01} & 4\pi e P_2 & -4\pi e\rho_{12} & -4\pi e\rho_{02} \\ 4\pi e\rho_{02} & 0 & -4\pi e\rho_{12} & 4\pi e P_1 & -4\pi e\rho_{01} \\ 0 & 0 & -4\pi e\rho_{02} & -4\pi e\rho_{01} & 4\pi e P_0 \end{bmatrix} \tag{3.22}$$

To prove the convexity of the IC function, it suffises to show that the hessian matrix is positive semi-definite. One of the tests to show this is to find the eigenvalues of the above matrix and see their signs.

To analyse the above matrix, let us consider a simpler case:

Case 1 - $X_0 = X_1$: In that case, $P_1 = P_0$ and $\rho_{01} = (P_0P_1)^{(1/2)} = P_0$. Thus the IC becomes a function of only 4 parameters, one of which is specified by the problem. The hessian matrix thus reduces to 3×3 matrix and can be expressed as follows:

$$\mathcal{H}_R(IC) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 4\pi e P_0 & -4\pi e P_0 \\ 0 & -4\pi e P_0 & 4\pi e P_0 \end{bmatrix} \tag{3.23}$$

If the above matrix has nonnegative eigenvalues then the matrix is positive semi-definite. For $\mathcal{H}_R(IC)$, the eigenvalues are $\{8\pi e P_0, 0\}$ which are both non-negative. Thus the hessian of the function is positive semi-definite and thus the Information constraint is convex in this case.

3.5.2 Marginal of a Multvariate Burr distribution

Consider a bi-variate Burr distribution with the following pdf:

$$f(x_1, x_2) = \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} c_1 d_1 c_2 d_2 \frac{x_1^{c_1-1} x_2^{c_2-1}}{(1 + d_1 x_1^{c_1} + d_2 x_2^{c_2})^{\alpha+2}} \tag{3.24}$$

where $\alpha, c_i, d_i, x_i > 0$, $i \in \{1, 2\}$

Let us now consider marginalising w.r.t. x_2 and see what we obtain:

$$\int_{x_2} f(x_1, x_2) dx_2 = \int_{x_2} \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} c_1 d_1 c_2 d_2 \frac{x_1^{c_1-1} x_2^{c_2-1}}{(1 + d_1 x_1^{c_1} + d_2 x_2^{c_2})^{\alpha+2}} dx_2 \tag{3.25}$$

$$= \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} c_1 d_1 c_2 d_2 x_1^{c_1-1} \int_{x_2} \frac{x_2^{c_2-1}}{(1 + d_1 x_1^{c_1} + d_2 x_2^{c_2})^{\alpha+2}} dx_2 \tag{3.26}$$

To solve the integral, consider the substitution $y_2 = x_2^{c_2}$. Thus $dy_2 = c_2 x_2^{c_2-1} dx_2$. The expression thus simplifies to:

$$
\begin{aligned}
&\frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} c_1 d_1 c_2 d_2 x_1^{c_1-1} \int_{y_2=0}^{\infty} \frac{x_2^{c_2-1}}{(1+d_1 x_1^{c_1}+d_2 y_2)^{\alpha+2}} \frac{1}{c_2 x_2^{c_2-1}} dy_2 && (3.27)\\
=&\frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} c_1 d_1 d_2 x_1^{c_1-1} \int_{y_2=0}^{\infty} \frac{1}{(1+d_1 x_1^{c_1}+d_2 y_2)^{\alpha+2}} dy_2 && (3.28)\\
=&\frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \frac{c_1 d_1 x_1^{c_1-1}}{d_2^{\alpha+1}} \int_{y_2=0}^{\infty} \frac{1}{(\frac{1+d_1 x_1^{c_1}}{d_2}+y_2)^{\alpha+2}} dy_2 && (3.29)\\
=&\frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \frac{c_1 d_1 x_1^{c_1-1}}{d_2^{\alpha+1}} \left(-\frac{1}{\alpha+1} \frac{1}{(\frac{1+d_1 x_1^{c_1}}{d_2}+y_2)^{\alpha+1}}\right)\Bigg|_0^{\infty} && (3.30)\\
=&\frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \frac{c_1 d_1 x_1^{c_1-1}}{d_2^{\alpha+1}} \left(\frac{1}{(\alpha+1)} \frac{d_2^{\alpha+1}}{1+d_1 x_1^{c_1}}\right) && (3.31)\\
=&\frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \frac{c_1 d_1 x_1^{c_1-1}}{1+d_1 x_1^{c_1}} = f(x_1) && (3.32)
\end{aligned}
$$

Thus we can see that marginalisation of bivariate Burr distribution w.r.t. to one random variable gives back the reduced Burr distribution for 1 variable, with the same parameters as the definition of the bivariate Burr distribution. This can easily be generalised for multivariable Burr distributions.

3.5.3 Proof of Proposition 6

Using the expression for entropy of a multivariate Burr distribution [38], we can write the information constraint for the case when $f(X_0, X_1, X_2)$ follow a joint multivariate burr distribution.

$$
\begin{aligned}
&H(X_0) + H(X_2) - H(X_0, X_1, X_2) \\
=& - log(\alpha) + (\alpha+1)\{\psi(\alpha+1) - \psi(\alpha)\} - logc_0 d_0^{1/c_0} + \frac{c_0 - 1}{c_0}\{\psi(\alpha) - \psi(1)\} \\
& - log(\alpha) + (\alpha+1)\{\psi(\alpha+1) - \psi(\alpha)\} - logc_2 d_2^{1/c_2} + \frac{c_2 - 1}{c_2}\{\psi(\alpha) - \psi(1)\} \\
& - \Bigg(- \sum_{i=0}^{2}(log(\alpha+i)) + (\alpha+3)\{\psi(\alpha+3) - \psi(\alpha)\} - \sum_{i=0}^{2} logc_i d_i^{1/c_i} \\
& + \sum_{i=0}^{2} \frac{c_i - 1}{c_i}\{\psi(\alpha) - \psi(1)\} \Bigg) \\
=& log\frac{(\alpha+1)(\alpha+2)}{\alpha} + 2(\alpha+1)\{\psi(\alpha+1) - \psi(\alpha)\} - (\alpha+3)\{\psi(\alpha+3) - \psi(\alpha)\} \\
& + logc_1 d_1^{1/c_1} - \frac{c_1 - 1}{c_1}\{\psi(\alpha) - \psi(1)\} \\
=& log\frac{(\alpha+1)(\alpha+2)}{\alpha} + 2(\alpha+1)\frac{1}{\alpha} - (\alpha+3)(\frac{1}{\alpha} + \frac{1}{\alpha+1} + \frac{1}{\alpha+2}) + logc_1 d_1^{1/c_1} \\
& - \frac{c_1 - 1}{c_1}\{\psi(\alpha) - \psi(1)\} \\
=& log\frac{(\alpha+1)(\alpha+2)}{\alpha} - \frac{\alpha^3 + 7\alpha^2 + 10\alpha + 2}{\alpha(\alpha+1)(\alpha+2)} + logc_1 d_1^{1/c_1} - \frac{c_1 - 1}{c_1}\{\psi(\alpha) - \psi(1)\}
\end{aligned}
$$

3.5.4 Marginal of a Multvariate Exponential distribution

Consider a bi-variate Exponential distribution with the following pdf:

$$
f(x_1, x_2) = \frac{\frac{\alpha}{\theta_1} e^{\frac{x_1 - \lambda_1}{\theta_1}} \frac{\alpha+1}{\theta_2} e^{\frac{x_2 - \lambda_2}{\theta_2}}}{(e^{\frac{x_1 - \lambda_1}{\theta_1}} + e^{\frac{x_2 - \lambda_2}{\theta_2}} - 1)^{\alpha+2}} \tag{3.33}
$$

where $x_i > \lambda_i$, $\alpha, \theta_i > 0$, $i \in \{1, 2\}$
Let us now consider marginalising w.r.t. x_2 and see what we obtain:

$$
\int_{x_2} f(x_1, x_2) dx_2 = \frac{\alpha}{\theta_1} e^{\frac{x_1 - \lambda_1}{\theta_1}} \frac{\alpha+1}{\theta_2} \int_{\lambda_2}^{\infty} \frac{e^{\frac{x_2 - \lambda_2}{\theta_2}}}{(e^{\frac{x_1 - \lambda_1}{\theta_1}} + e^{\frac{x_2 - \lambda_2}{\theta_2}} - 1)^{\alpha+2}} dx_2 \tag{3.34}
$$

To solve the integral, consider the substitution $y = e^{\frac{x_2 - \lambda_2}{\theta_2}}$. Thus $dx_2 = \frac{\theta_2}{y} dy$. The integral thus simplifies to:

$$\frac{\alpha}{\theta_1}e^{\frac{x_1-\lambda_1}{\theta_1}}\frac{\alpha+1}{\theta_2}\int_1^{\infty}\frac{y}{(e^{\frac{x_1-\lambda_1}{\theta_1}}+y-1)^{\alpha+2}}\frac{\theta_2}{y}dy \tag{3.35}$$

$$=\frac{(\alpha)(\alpha+1)}{\theta_1}e^{\frac{x_1-\lambda_1}{\theta_1}}\int_1^{\infty}\frac{1}{(e^{\frac{x_1-\lambda_1}{\theta_1}}+y-1)^{\alpha+2}}dy \tag{3.36}$$

$$=\frac{(\alpha)(\alpha+1)}{\theta_1}e^{\frac{x_1-\lambda_1}{\theta_1}}\int_1^{\infty}\frac{1}{(e^{\frac{x_1-\lambda_1}{\theta_1}}+y-1)^{\alpha+2}}dy \tag{3.37}$$

$$=\frac{(\alpha)(\alpha+1)}{\theta_1}e^{\frac{x_1-\lambda_1}{\theta_1}}\frac{-1}{\alpha+1}\left((e^{\frac{x_1-\lambda_1}{\theta_1}}+y-1)^{-(\alpha+1)}\right)\Big|_1^{\infty} \tag{3.38}$$

$$=-\frac{(\alpha)}{\theta_1}e^{\frac{x_1-\lambda_1}{\theta_1}}\left((e^{\frac{x_1-\lambda_1}{\theta_1}}+y-1)^{-(\alpha+1)}\right)\Big|_1^{\infty} \tag{3.39}$$

Note that $lim_{y\to\infty}(e^{\frac{x_1-\lambda_1}{\theta_1}}+y-1)^{-(\alpha+1)}=0$ since $\alpha>0$. Thus, we get :

$$-\frac{(\alpha)}{\theta_1}e^{\frac{x_1-\lambda_1}{\theta_1}}(-((e^{\frac{x_1-\lambda_1}{\theta_1}}+1-1)^{-(\alpha+1)})) \tag{3.40}$$

$$=\frac{(\alpha)}{\theta_1}e^{\frac{x_1-\lambda_1}{\theta_1}}(e^{\frac{x_1-\lambda_1}{\theta_1}})^{-(\alpha+1)}=f(x_1) \tag{3.41}$$

3.5.5 Proof of Proposition 7

Using the expression for entropy of a multivariate exponential distribution [38], we can write the information constraint for the case when $f(X_0, X_1, X_2)$ follow a joint multivariate burr distribution.

$$\begin{aligned}
&H(X_0)+H(X_2)-H(X_0,X_1,X_2)\\
&=-log(\frac{\alpha}{\theta_0})+1-log(\frac{\alpha}{\theta_2})+1-\Bigg(-(log(\frac{\alpha}{\theta_0})+log(\frac{\alpha+1}{\theta_1})+log(\frac{\alpha+2}{\theta_2})\\
&\qquad+(\alpha+3)(\frac{1}{\alpha}+\frac{1}{\alpha+1}+\frac{1}{\alpha+2}-\frac{3}{\alpha}\Bigg)\\
&=log\frac{(\alpha+1)(\alpha+2)}{\alpha\theta_1}-\frac{(\alpha^2+7\alpha+7)}{(\alpha+1)(\alpha+2)}
\end{aligned}$$

Part II

Causal Information Structure

Chapter 4

Information theoretical bounds

To know that
we know what we know,
and to know that
we do not know what we do not know,
that is true knowledge.

Nicolaus Copernicus

4.1 Introduction

In this part we turn our attention to causal information structure. Compared to the previous part, the essential difference is that in this part, no agent is assumed to have a non-causal knowledge of the nature state. In other words, no agent knows what the future holds in store. This assumption is more applicable to real world scenarios as one would be hard-pressed to find applications where the future realizations of the nature state is known with reasonable accuracy to some agent.

Not knowing the future realizations removes the information assymetry inherent in the previous part. This has the following ramnifications :

1. While the previous part dealed entirely with coding through actions, also known as implicit communication in the literature, communication between agents does not figure in this part (apart from the application to Witsenhausen Counterexample). This is simply because no agent has any information about the future realizations of the nature state to communicate. In the case where an agent cannot even observe the present realization of the nature state, communication through actions indeed plays a significant role. We briefly discuss this scenario in the application to Witsenhausen Counterexample.

2. Also, throughtout the previous part, we showed that the performance limits of coordination schemes with non-causal information structures are characterized by information constraints on the implementable distrbutions. In this section, since there is no information transfer between the agents, information constraints do not appear in our analysis.

3. In this part, due to the symmetry in the information structure, we consider co-ordination scenarios involving more than 2-Agent teams. This greatly enhances the applicability of the framework developed in this part.

While in the previous part, we contended ourselves with the performance limits of coordination schemes, without ever expliciting them, in this part, we shall attempt to provide practical coordinaton schemes for certain important technological challenges. To this effect, we exploit a recent theorem derived in [8] to find power control functions which may exploit the available knowledge optimally (in the long run). We devise a decentralized offline algorithm which requires knowledge of the statistics of the nature state X_0 as well as the noisy channels $\daleth(S_i|X_0)$ of the feedback for each agent. Based on the feedback available to each agent about the nature state, we are capable of devising coordination schemes for any application following the given information structure. We specifically apply our framework to the case of power control schemes in wireless communications for various objectives, as well as co-ordination schemes in Smart Grids to even out the consumption of electrcity during a day.

The key insight in the proposal is that most decentralized decision schemes proposed in literature hinge on game theoretic frameworks, which does not necessarily attemmt to achieve social optimality. This is because co-operation is not assumed while considering Nash Equilibria, and the schemes proposed, though robust and fair, do not use the information avialable to agents to coordinate optimally. Our attempt is to get as close as possible to using the knowledge available at each agent in determining optimal decision functions maximizing co-ordination and thus sum-performance.

In this chapter, we develop this general framework, together with the proposed algorithm to devise practical schemes. Indeed, one of the salient features of the proposed framework is to provide robust decision functions by accounting for noise by incorporating noise statistics while determining the decision functions. In Sec. 4.2, we recall the theorem from [8] which characterizes the implementable distributions under the information structure considered. We exploit this theorem to devise a sub-optimal algorithm for finding decision functions in Sec. 4.3.

4.2 Performance limits of an N-Agent Coordination problem

Consider a coordination problem where N Agents are trying to coordinate their decision $X_i \in \mathcal{X}_i$, $i \in \mathcal{N} = \{1, ..., N\}$. The coordination is done with respect to a nature state represented by $X_0 \in \mathcal{X}_0$ with the goal of optimizing a common payoff function $w(x_0, x_1, x_2, ...x_n)$ over a long time period T. The nature state X_0 for applications to wireless communications, for example, could be channel gain coefficients $g_{ij} = |h_{ij}|^2$, with g_{ij} being the channel gain of the link between transmitter i and receiver j. X_0 denotes a random nature state which affects the common payoff function for the system and is not controlled by the coordinating agents. The realizations of the nature state $X_{0,t}$ at each time instant t are i.i.d. and follow a probability distribution ρ. In wireless communications ρ is typically an exponential distribution for each channel gain g_{ij}. Our aim is to characterize all the achievable expected payoffs under this certain information structure over a long time period T, $T \to \infty$.

Firstly, we need to formally define the information structure under consideration. At every instant t, agent i is assumed to have an image or a partial observation $S_{i,t}$ of the nature state $X_{0,t}$ with respect to which all agents are coordinating. Again, for wireless communications, this could be the knowledge of local Channel State Information at Transmitter (CSIT), i.e. transmitter i observes a noisy version (in general) of only the direct link channel gain g_{ii}. One could imagine other kinds of information available at the transmitters: for example, transmitter i observes all the links g_{ji}, $\forall j$. The observations $S_{i,t}$ are assumed to be generated by a memoryless channel whose transition probability is denoted by $\daleth_i(S_{i,t}|X_{0,t})$. All agents have to make their decision $X_{i,t}$ based on this information received. Formally, the sequence of decision functions for agent i, $f_{i,t}$, is defined as:

$$\begin{aligned} f_{i,t} \quad : \quad & \mathcal{S}_i^t \times \mathcal{U} && \longrightarrow \quad \mathcal{X}_i && (4.1)\\ & (s_i(1), s_i(2)...s_i(t), u(t)) && \longmapsto \quad x_i(t) && (4.2) \end{aligned}$$

where $\mathcal{S}_i^t = \mathcal{S}_i(1) \times \mathcal{S}_i(2)... \times \mathcal{S}_i(t)$ is the discrete observation alphabet till the instant t and $\mathcal{X}_i$ is the action chosen by agent i with $|\mathcal{S}_i|, |\mathcal{X}_i| < \infty$. $\mathcal{U}$ is the alphabet of the auxiliary variable U which is discussed in more detail later. $s_i(t)$, $u(t)$ and $x_i(t)$ are the realizations of the corresponding variables at instant t.

The problem is said to be decentralized as each agent chooses its action independently based on the information received by it. Since the nature state is an external random variable, and thus not controled by agents, the quantity to be optimized is the expected objective function:

$$\mathbb{E}_Q[w(X)] = \sum_{x \in \mathcal{X}} w(x_0, x_1, ..., x_N) Q(x_0, x_1, ..., x_N) \tag{4.3}$$

where $X = (X_0, X_1, ..., X_N)$, $x = (x_0, x_1, ..., x_N)$, and $\mathcal{X} = \mathcal{X}_0 \times \prod_{i=1}^N \mathcal{X}_i$. $Q(x_0, x_1, ..., x_N)$ is the joint probability distribution of the variables affecting the payoff. An important point to note is that since expectation is a linear operator, optimizing the expected payoff is equivalent to finding the optimal distribution $Q(x_0, x_1, ..., x_N)$. However, the optimization problem is not so trivial as indeed there are certain restrictions on the distributions Q that are implementable given the imposed information structure. We now define the notion of an implementable distribution.

Definition 6 (Implementability). *Let the information structure be as defined in (4.1). The probability distribution $\overline{Q}(x_0, x_1, ..., x_N)$ is implementable if there exist decision functions $\sigma_{i,t}$ such that as $T \to +\infty$, we have for all $x \in \mathcal{X}$,*

$$\frac{1}{T}\sum_{t=1}^{T} Q_{X_0 \ldots X_N, t}(x_0, ..., x_N) \longrightarrow \overline{Q}(x_0, ..., x_N) \tag{4.4}$$

where $Q_{X_0 X_1 \ldots X_N, t} = Q_{X_1, \ldots X_N | X_0, i} \times \rho$ is the joint distribution induced by $\sigma_{i,t}$ at stage t.

As seen before, the expected payoff is characterized by the probability distribution Q over all the variables that intervene in the payoff function. Thus, the time averaged expected payoff $\overline{w}$ is said to be achievable, if and only if the corresponding distribution $\overline{Q}$ is implementable. The following theorem characterizes the achievable payoffs that are implementable under the information structure (4.1).

Theorem 4. *[8] Assume the random process $X_{0,t}$ to be i.i.d. following a probability distribution ρ and the available information to the transmitters $S_{i,t}$ to be the output of a discrete memoryless channel obtained by marginalizing the conditional probability $\daleth(s_1, ..., s_N | x_0)$. An expected payoff $\overline{w}$ is achievable in the limit $T \to \infty$ if and only if it can be written as:*

$$\begin{aligned} \overline{w} = \sum_{\substack{x_0, x_1, \ldots x_N, \\ u, s_1, \ldots s_N}} & \rho(x_0) P_u(u) \daleth(s_1, \ldots s_N | x_0) \times \\ & \left(\prod_{i=1}^{N} P_{X_i | S_i, U}(x_i | s_i, u)\right) w(x_0, x_1, \ldots x_N) \end{aligned} \tag{4.5}$$

where U is an auxiliary variable which can be optimized and $P_{X_i|S_i,U}(x_i|s_i, u)$ is the probability that Transmitter i, chooses action x_i after observing s_i, u.

The auxiliary variable U is an external lottery known to the transmitters beforehand, which can in general, be used to achieve better coordination. However, we do not exploit this auxiliary variable in our analysis, and thus we shall omit it in the following analysis for greater clarity.

While theorem 4 provides us with all the achievable payoffs given the information structure defined in equation (4.1), it does not provide optimal sequences of decision functions $f_{i,t}$ that might help achieve this payoff. One of the main aims of this chapter is precisely to provide a procedure to obtain those decision functions with reasonable complexity. To achieve this, certain simplifications and observations need to be made. Firstly, instead of looking for a sequence of functions $f_{i,t}$, we shall only search for optimal stationary strategies, i.e. $f_i = f_{i,t} \forall t$. This induces no performance loss as the nature state is considered to be i.i.d. for every timeslot, and thus the optimal strategies are also time-independent.

Another important observation is that theorem 4 characterizes all achievable payoffs in terms of conditional probabilities $(P_{X_1|S_1}, ..., P_{X_N|S_N})$. However, if one considers the optimal achievable payoff, the conditional probabilities $(P_{X_i|S_i}$ simplify to mere functions $f_i(s_i)$ due to proposition 8.

Proposition 8. *Restricting non deterministic conditional probability $P_{X_i|S_i}$ in theorem 4 to deterministic functions $f_i(s_i)$ induces no loss in optimality*

Proof. Assume that at the optimal point, there exists one agent (agent i) with actions, labelled X_i^1, X_i^2, with non zero probabilities. If it were the case, the agent could always increase the common team payoff by choosing the action which gives better performance with probability 1; thus leading to the conditional probability $P_{X_i|S_i}$ being a deterministic function $f_i(s_i)$. This is a classical argument to show that the solution of a multilinear program lies at the vertices.[39] □

where the decision functions of interest, f_i, are formally defined as follows:

$$\begin{array}{rcllcl} f_i & : & \mathcal{S}_i & & \longrightarrow & \mathcal{X}_i \\ & & s_i(t) & & \longmapsto & x_i(t) \end{array}$$

With these simplifications, we can now restate the theorem 4 for optimal achievable payoff $\overline{w}^\star$ in terms of the optimal decision functions $f_i^\star$.

$$\overline{w}^\star = \sum_{x_0, s_1, \dots s_N} \rho(x_0) \daleth(s_1, \dots s_N | x_0) w(x_0, f_1^\star(s_1), \dots, f_N^\star(s_N)) \tag{4.6}$$

Note that in terms of optimality, all the simplifications made till now induce no performance loss. However, jointly finding the optimal functions $f_i^\star(s_i)$ still entails high complexity. Indeed the complexity of joint optimization could be prohibitive if the number of agents N or alphabet size $|\mathcal{X}_i|, i \in 0, \dots, N$ is too high. In the following section, we endeavour to devise a distributed algorithm, which although sub-optimal, is tractable complexity-wise.

4.3 Algorithm for finding sub-optimal decision functions

To reduce the complexity of the optimization problem further, we use the sub-optimal approach of sequential best response dynamics [40], with each user sequentially updating their decision function to maximize the common payoff. This approach consists of each transmitter choosing the best decision function, keeping the other decision functions constant, and each transmitter doing so sequentially within an iteration. This procedure is is repeated till convergence is achieved, typically in a few iterations. The number of iterations required for convergence of course depends on the number of agents coordinating, but it scales up very slowly as algorithms based on best-response dynamics normally converges very fast [40]. Furthermore, since we are considering a common 'team' payoff $w(x_0, x_1, \dots, x_N)$, the convergence of best-response dynamics is guaranteed (See 9). Note however that BRD procedures do not necessarily converge to globally optimum solutions. Nonetheless, the decision functions obtained by the simplification are typically good and easily outperform solutions which guarantee convergence to Nash Equilibria.

To apply best response dynamics, we rewrite the expected utility isolating the sum w.r.t. s_i, the observation of agent i.

$$\begin{aligned}\overline{w} &= \sum_{x_0, s_1, \ldots s_N} \rho(x_0) \urcorner(s_1, \ldots s_N | x_0) w(x_0, f_1(s_1), \ldots, f_N(s_N)) \\ &= \sum_{x_0, s_i} \rho(x_0) \urcorner(s_i | x_0) \sum_{s_{-i}} \urcorner(s_{-i} | x_0) w(x_0, f_1(s_1), \ldots, f_N(s_N))\end{aligned}$$

where s_{-i} represents the vector comprising of observations of the agents other than agent i.

To update the function $f_i(s_i)$ for agent i, all we need is an action corresponding to each possible observation s_i. Thus, for every fixed s_i, we need to find the action x_i which maximizes the expected payoff. More formally, we define the following quantity

$$\theta(s_i, x_i) = \sum_{x_0} \rho(x_0) \urcorner(s_i | x_0) \sum_{s_{-i}} \urcorner(s_{-i} | x_0) w(x_0, x_i, f_{-i}(s_{-i})) \tag{4.7}$$

In essence, this is the expected common utility w.r.t. the nature state x_O, if agent i choses the action x_i assuming that it knows the decision functions of all other agents $f_{-i}(s_{-i})$. The update of the function f_i for the observation s_i is then simply $\arg\max_{x_i} \theta(s_i, x_i)$. Updating the function for each possible observation completes the update for an agent, and this procedure is repeated sequentially for other agents. The algorithm is said to have converged if after an iteration, i.e. one round of function updates for all agents, the difference in the functions from the previous iteration and the updated functions is below the tolerance level.

An important point to note is that the transmitters can run this algorithm offline as all they need to know are the channel statistics ρ, and the channel $\urcorner$. This helps in exploiting the decision functions for the entire timeslot during which the nature state remain constant. Thus we exploit the channels from the very start of a timeslot as opposed to taking some time initially to find the optimal functions online. The latter case is considered in some decentralized solutions such as the algorithms based on water-filling techniques proposed in [41], [42]. It is easy to see that this algorithm necessarily converges for a common payoff, as stated in 9

Proposition 9. *Algorithm 1 converges for a common performance criteria* $w(x_0, x_1, ..., x_N)$.

Proof. The result can be proved by calling for an exact potential game property [43] (the argument may hold in the more general case in which the transmitter have different performance criteria). □

Algorithm 1: Proposed decentralized Algorithm for finding decision functions for all agents

inputs : $\mathcal{X}_i$, $\forall i \in \{0, ..., N\}$, $w(x_0, x_1, ...x_N)$, $\forall \underline{x}$,
$\rho(x_0)$, $\daleth(\underline{s}|x_0)$, $\forall x_0$, f_i^{init}, $\forall i \in \{1...N\}$
output: $f_i^*(s_i, u_i)$, $\forall i \in \{1...N\}$

Initialization: $f_i^0 = f_i^{\text{init}}$, $iter = 0$, $iter_{max} = 100$

while $\exists i : f_i^{\text{iter}-1} - f_i^{\text{iter}} \geq \epsilon$ AND $iter \leq iter_{max}$ OR $\text{iter} = 0$ **do**
    $\text{iter} = \text{iter} + 1$;
    foreach $i \in \{1, \dots, N\}$ **do**
        foreach $s_i \in \mathcal{S}_i$ **do**
            $f_i^{\text{iter}}(s_i) = \arg\max_{x_i} \theta(s_i, x_i)$ using (4.7);
        end
    end
end

With these tools in hand, we are ready to tackle some practical co-ordination problems from various domains. While the simplifications made to reduce the complexity might induce some performance laws, we shall see in the next chapter that it still performs better than the decentralized solutions currently proposed by the literature.

Chapter 5

Application to N-Agent team coordination problems

> The hallmark of a good theory is that with hindsight, it is evident.
>
> Anonymous

In this chapter, we shall consider few specific applications of the theory developed in Chap. 4. In particular, we shall consider 3 applications: 1) Power Control in parallel interference channel for wireless communications 2) Power consumption scheduling for applications in Smart Grids 3) Finding decision functions for the famous Witsenhausen Counterexample. The applications are mutually independent, even though they all rely on the same underlying framework developed in the previous chapter.

5.1 Power Control in Wireless Communications

As an application, we treat the well studied problem of optimal power control in interference channels ([44], [45], [46]) and show that not only does our framework attain various payoffs better or identical to the state-of-the-art, it also provides the optimal power control functions in diverse scenarios. The scenarios considered resemble those of [41], [42]. These works propose ad-hoc solutions for specific utilities but not a generic framework which directly provides a power control function that aims at exploiting the available and arbitrary information as optimally as possible. Indeed our approach is very general and can be used to treat more complex and interesting scenarios, e.g. by considering different common objectives, robust power control taking the noisy communications into account, vectorial optimization for power allocation schemes etc.

Eventhough the framework developed previously is very general and conducive to optimizing different utilities such as sum-rate, sum-throughput, sum-energy [47], we develop the utility sum-energy utility in detail. This choice is motivated by the increasing importance of energy efficiency for wireless communications, as well as the difficulty for finding appropriate solutions for it. For example, [48] showed that binary power control is almost optimal for maximizing sum-rate. [49] later showed that thresholding strategy is optimal for sum-rate utilities if only binary power control policies are considered. We generalize those results for the case of energy-efficiency by proposing a combination of thresholding strategy and channel inverson policy, and show its full relevance in the problem of optimal power control for energy efficient communications.

Indeed, for the case of power control in wireless interference channels, these assumptions have been paid some attention. The necessity of minimizing the information required at the transmitters for coordination led to approaches proposed in [49], [48], [50]. They consider only local channel state information at the transmitters (CSIT). This is also important for reducing the complexity of the optimizations to be performed. For sum-rate maximization, the scenario considered here is comparable to a single-carrier version of the iterative water filling algorithm (IWFA) [51]. In the case of sum-energy utility, channel inversion strategy was proposed by [1], whereas the multi-carrier version of the problem was considered by [52]. Decentralization of the power control, or for that matter for any wireless network design, is the reason why game theory has been extensively applied to such problems [40]. We thus compare our results to game theoretic equilibria analysis as most of the decentralized optimization literature only guarantees convergence to Nash equilibria.

We structure this section as follows: In 5.1.1, we describe the system model of the power control problem in a parallel interference channel. We then proceed to provide simulation results in 5.1.2 with detailed analysis about various aspects of the algorithm proposed in the case of sum-energy utility.

5.1.1 System Model

Consider N single-antenna Transmitter-Reciever pairs Tx_i, Rx_i, $i \in \{1, 2..., N\}$ communicating over a single-carrier parallel interference channel with the channel coefficients being $g_{ij} \in \mathcal{G}_{ij}$, $i, j \in \{1, 2..., N\}$ being the index for Transmitter i and Receiver j respectively. The channel coefficients could take values in the *discrete* alphabet $\mathcal{G}_{ij}$. The transmitter i transmits at powers $P_i \in \mathcal{P}_i$. Due to the power constraint at every transmitter, the total power consumed by transmitter i should not exceed P_{max}, i.e. $P_i \in [0, P_{max}]$. The utility u_i for the pair Tx_i, Rx_i, is typically a function of SINR $\gamma_i = \dfrac{P_i g_{ii}}{\sigma^2 + \sum\limits_{i \neq j} P_j g_{ji}}$. Without loss of generality, we shall consider $\sigma^2 = 1$ henceforth. The SNR P_{max}/σ^2 will be thus regulated by varying P_{max}

A well accepted model of statistics for the channel gains $g_{ij} = |h_{ij}|^2$ is Rayleigh fading. In this model, due to central limit theorem, the real and the imaginary components of h_{ij} follow a normal distribution, and thus the channel gain g_{ij} follows exponential distribution ($|h_{ij}|$ follows a Rayleigh distribution).

We also assume that all channel gain distributions are independent of each other and that their realizations are i.i.d.

We shall consider the following common utilities for our analysis.

- Sum-rate: $w_R(\underline{P}, \underline{g}) = \sum_{i \in \mathcal{N}} \log(1 + \gamma_i)$;
- Sum-goodput: $w_G(\underline{P}, \underline{g}) = \sum_{i \in \mathcal{N}} \Omega(\gamma_i)$;
- Sum-energy: $w_E(\underline{P}, \underline{g}) = \sum_{i \in \mathcal{N}} \dfrac{\Omega(\gamma_i)}{P_i}$.

Typical functions for Ω are $\Omega(\gamma) = e^{-\frac{c}{\gamma}}$ [53], where $c > 0$ or $\Omega(\gamma) = (1 - e^{-\gamma})^M$ where $M \geq 1$ is the packet length [1]. In simulations, we chose the former function and further supposed $c = 1$ as it only changes the optimal solutions by a multiplicative constant. Note that Proposition **??** holds for the above payoffs as we are considering sums of individual payoffs which trivially satisfy the required condition for convergence.

The following table explicits the relation between the input variables for Algo. 1 and the corresponding quantities for the application to power control defined above.

General model	**Power Control Application**
Nature State - X_0	Channel state - $(g_{11}, ..., g_{NN})$
Decision of Transmitter i - X_i	Power emitted $P_i \in \mathcal{P}_i$
$w(x_0, x_1 ... x_N)$	$w_R(\underline{P}, \underline{g}), w_G(\underline{P}, \underline{g}), w_E(\underline{P}, \underline{g})$
$\rho(x_0)$	$\prod_{i=1}^{N} \prod_{j=1}^{N} e^{-\overline{g}_{ij}}$, $\overline{g}_{ij} = \mathbb{E}(g_{ij})$
$\daleth(s_i \vert x_0)$	$s_i = g_{ii}$ or $s_i = \widehat{g_{ii}}$
f_i^{init}	$f_i(s) = P_{max}$, $\forall s \in \mathcal{S}$

Table 5.1: Correspondence between the general framework and its application to power control

5.1.2 Numerical Results

Simulation Setup

For $\mathcal{G}_{ij}$, since the channel coefficients g_{ij} follow exponential distribution, a uniform discrete set is typically not a good representative due to varying probabilities for every representative. In our simulations, for $|\mathcal{G}_{ij}| = n$, we find the representative points such that the interval corresponding to each point has an equal probability $1/n$. However, the size of alphabets $|\mathcal{G}_{ij}| = n, \forall i, j \in \{1, ..., K\}$ is kept same for all channel gains. The average value for the direct channel gains $\mu_{ii} = 1$, and the interfering channel gains $\mu_{ij} = 0.3, i \neq j$, unless otherwise stated. The power levels $\mathcal{P}_i$ are quantized uniformly in dB in the interval $[-30dB, 10dB]$.

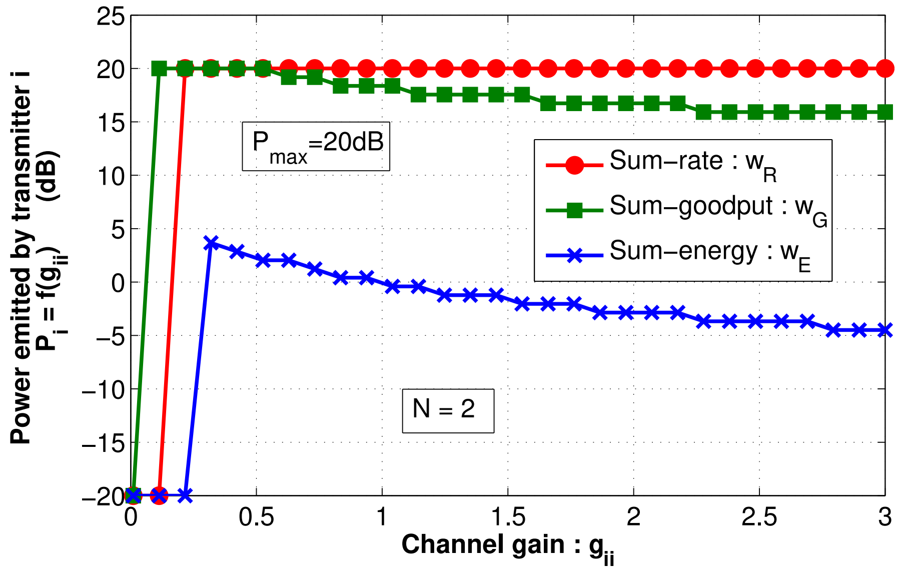


Figure 5.1: The figure represents the power control functions f_i provided by Algo. 1 for the three performance metrics under consideration for power control problem. In particular, it is seen that for maximizing sum-energy, the obtained power control function exhibits a threshold under which transmission should not occur for a given transmitter.

While considering no noise in the feedback received at transmitter i, the observation alphabet is considered to be the same a

Features of the Algorithm for applications in Wireless communications

Here, we shall illustrate the salient features of the proposed algorithm for power control applications. To illustrate the features, we will concentrate on the utility sum-energy as it is the most novel and important result found using the algorithm.

- *Power Control for different utilities*: In Fig. 5.1, we plot the power control functions f_i obtained using Algo. 1 for the case $N = 2$ against g_{ii} the direct channel gain between the i^{th} Transmitter-Receiver pair. We do not consider noise in the channel estimation for the moment, and thus in this case $s_i = g_{ii}$. Three different utilities are considered, sum-rate, sum-energy, and sum-goodput.

  For the case of sum-rate, it is known that binary power control $P_i \in \{P_{min}, P_{max}\}$ is optimal for 2 Transmitter-Receiver pairs [48]. Moreover, as shown in [49], optimal power control functions with only local CSIT

(g_{ii}) amounts to $P_i(g_{ii}) = P_{min}$ if $g_{ii} \leq g^*$ and $P_i(g_{ii}) = P_{max}$ otherwise. It is reassuring to find that our results verify this. Thus, in the case of sum-rate with local CSIT, we find exactly the same results as state-of-the-art.

In the case of sum-energy, we see that all the available power P_{max} is not used. Indeed, even in the case of only one Transmitter-Receiver pair, the optimal power control function is c/g_{11}. We see that there is also a threshold value of g_{ii} below which, $P_i = P_{min}$. Above the threshold value, the function is similar to the optimal solution obtained in the case of only one Transmitter-Receiver pair. The threshold function is also seen as solution in the case of sum-packet rate. This is because the utility function is not monotonous w.r.t. SINR. To the best of our knowledge, the power control functions found for sum-energy are new. Solutions in literature do not propose thresholding of power control functions to help reduce interference in case of bad direct channel gains.

- *Performance gain w.r.t. Nash Equilibria*: As stated before, since our framework is co-operative, it easily outperforms game theoretical solutions which guarantee convergence to Nash equlibria under certain conditions. In Fig. 5.2, we compare the performance of the power control function with that of Nash equilibrium for different SNRs in the case of sum-energy and sum-rate. Thus, on the $y - axis$ we see the relative performance gain ($w_f - w_{Nash}/w_{Nash}$ in %) of our algorithm when compared to average payoffs obtained by Nash equilibrium. Current state-of-the-art proposes solutions which converge to the Nash equilibrium. Thus, the solutions are generally not optimal while maximizing common utilities that are sum of the individual utilities. The evaluations for obtaining average expected payoffs was done for 10^6 channel realizations.

- *Different feedback*: Another salient feature of our framework is its ability to incorporate different kinds of feedbacks available at transmitters. This helps in analyzing various possible feedbacks which could help improve the performance. However, the possible additional information at transmitter does come at a cost, and we do not analyze the tradeoff involving this communication cost and the performance gain. Also, we see that additional feedback does not bring significant performance gain.

  The more informed choice evidently brings some gain in payoff, as seen from Fig. 5.3. However, suprisingly we see that there is very little difference between the feedback being the vector $g_{ji} \forall j \in \{1, ..., N\}$ and just the direct channel gain g_{ii}. This could be explained by the fact that the decision function $f_j(g_{1j}, ...g_{Nj})$ for user $j, j \neq i$ does not depend on g_{ji}. Thus the additional information does not help in better coordination and we obtain similar payoffs.

  We also see that distributed decision making knowing the global CSI $g_{ij} \forall i, j \in \{1, ..., N\}$ performs almost at centralized social optimum. Centralized social optimum is found through joint exhaustive search for the

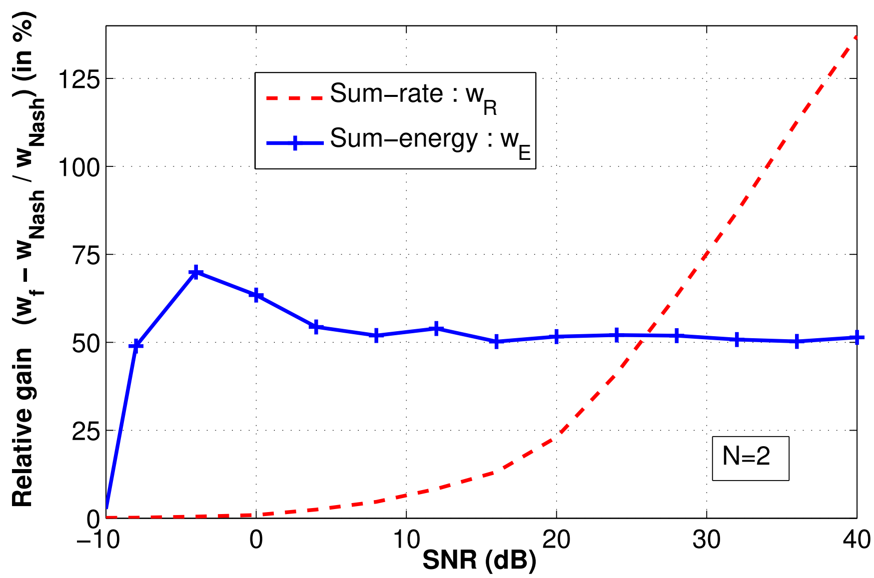


Figure 5.2: Relative performance gain ($w_f - w_{Nash}/w_{Nash}$ in %) of our algorithm when compared to average payoffs obtained by Nash equilibrium for sum-rate and sum-energy. The curve for sum-energy saturates as at high SNR, P_{max} is not utilized as power emitted is much lower.

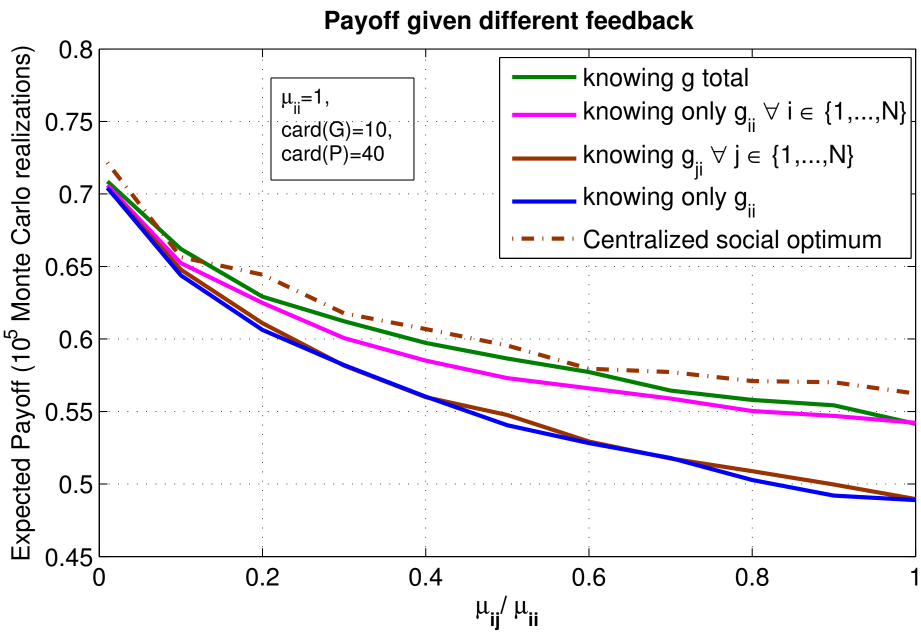


Figure 5.3: Evaluation of different information available at transmitters for different levels of interference. Most importantly, one notices that transmitter knowing the interference channel gains g_{ij},$i \neq j$ does not bring any performance gain.

optimal power levels for every realization of channel coefficients. This is again an indication of the 'goodness' of our proposed distributed algorithm.

For a loose lower bound, we consider the case of no feedback to the users. In the absence of channel feedback but knowing the channel statistics, each user naively chooses the power level which maximises their utility on an average assuming no interference. This power level , in the case of energy efficiency u_{EE}, corresponds to $1/\mu_{ii}$ which is the mean of the direct channel coefficient g_{ii}.

- *Robustness to Noisy Feedback* : One of the advantages of our framework is that it provides optimal power control functions even for noisy channel estimates. We illustrate this in Fig. 5.4 where power control functions for different levels of noise in channel estimation are plotted. This noise simulates the error in estimation due to noise during feedback transmission, or just simply estimation error of the channel state at the receiver. The noise for the simulations is gaussian, i.e. $\hat{g}_{ii} = g_{ii} + z$ where $Z \sim \mathcal{N}(0, \sigma_z^2)$, with $\sigma_z \in \{0, 1, 3\}$. Notice that since the alphabet $\mathcal{G}_{ii}$ is discrete, whereas the noise is continuous, we requantize the noisy feedback with the same

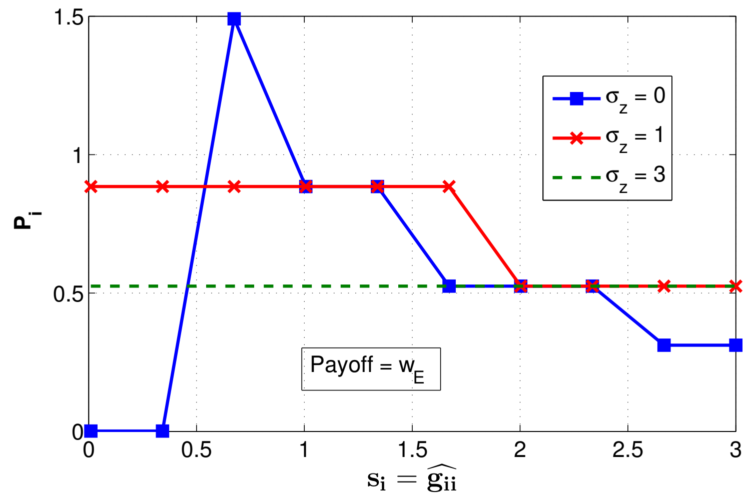


Figure 5.4: Influence of channel estimation noise on the power control functions. We see that as noise increases, the optimal power emitted becomes more uniform. This is intuitive as higher the noise, higher the uncertainty in the observation leading to same power emitted for all channel gains observed in the asymptotic case.

alphabet i.e. $\mathcal{S}_i = \mathcal{G}_{ii}$.

As expected, the power control functions become more uniform at higher noise levels, as the information received is less reliable, and thus transmitters emit at a power level which maximizes the utility after averaging over the uncertainty in observation due to the noise.

- *Robustness to noisy estimation of channel statistics* : So far, we have assumed that the channel statistics $\mu_{ij} \forall i, j \in \{1, ..., K\}$ are known perfectly to everyone. In reality, this might not be the case due to many reasons: error in estimation of the channel, added noise during the communication of statistics being a few. To see the influence of this error, we run our algorithm intentionally using erroneous statistics $\widehat{\mu_{ij}} = \mu_{ij}(1 + Z)$. Here Z is akin to percentage error in the estimation of channel statistics, with $z = n$ corresponding to $100n\%$ error in estimation.

  We see form the Fig. 5.5, the average payoff drops off only slightly for positive Z, i.e. if one overestimates the channels. However, interestingly, if one underestimates the channels, the average payoffs are considerably less. This might seem strange at first glance, but closer inspection of the power control functions for the different cases clarifies the reason for this phenomenon.

  From Fig. 5.6, we see that the threshold value for g_{ii}, above which the

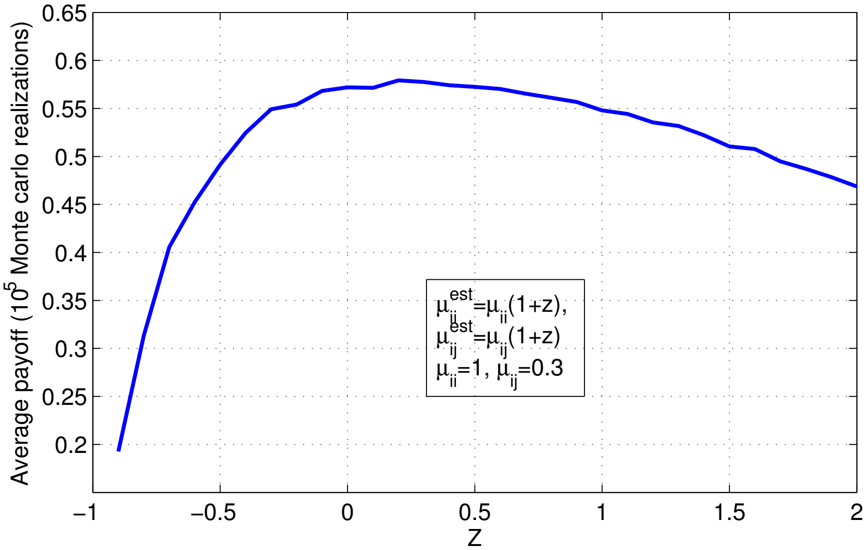


Figure 5.5: Influence of error in estimation of channel statistics on average payoff. We see that negative error in statistics (underestimation of channels) leads to worse payoffs than positive error.

transmitter emits, is lesser in the case of negative Z (underestimation). Moreover, since the chosen power levels are roughly equal to $1/g_{ii}$, transmitter emits at a very high power level creating high interference. On the other hand, in the case of positive Z, while the threshold is higher, the chosen power levels are not much different, and thus there is less loss in performance.This indicates that its better to overestimate the channel statistics than to underestimate it.

Analysis of the algorithm

- ***Discrete alphabets*** $\mathcal{G}_{ij}$, $\mathcal{P}_i$: Now we shall analyze the influence of discretization of feedback $\mathcal{S}_i$, as well as Power levels $\mathcal{P}_i$ on the average payoff obtained for our algorithm. Indeed, the complexity of the algorithm is sensitive to the size of these alphabets. Here, we show that the restrictions on the alphabets being discrete and small induce virtually no loss in optimality when compared to much finer discretizations. Furthermore, the small alphabet provides greater robustness against uncertainty in the observations and estimations; i.e. noise in feedback, and higher error in estimation of channel statistics.

  For the simple case where number of Tx-Rx pairs , $N = 2$, we did Monte-Carlo simulations for the functions found using the Algo. 1 while changing

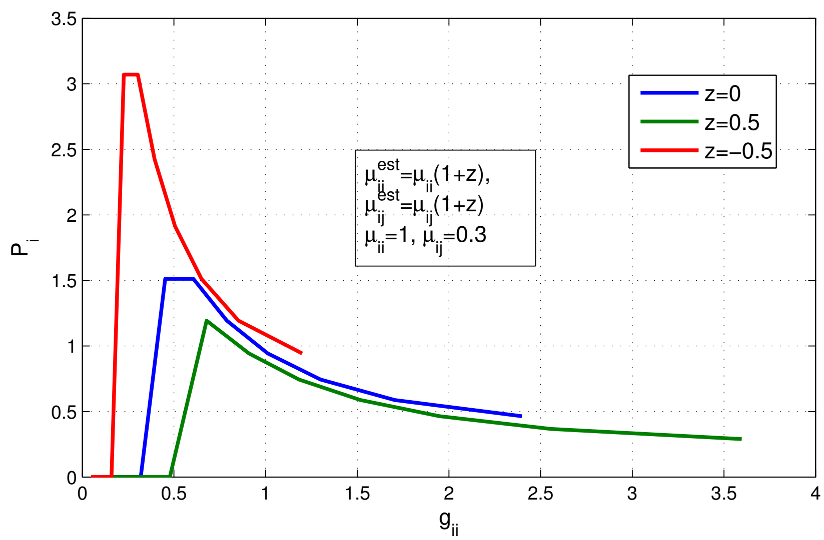


Figure 5.6: Power control functions for different positive and negative error in estimation of channel statistics. It provides an explanation for the greater payoff loss due to underestimation of channel statistics (negative z). We see that the threshold λ_i due to negative error is smaller. This results in higher power emited P_i which causes interference with other users.

the cardinality of $\mathcal{P}$ and $\mathcal{G}_{ij}$. The cardinality was changed so as to conserve the equal probability quantization, but also have a nested structure. This was done to see the true effects of finer discretization. Without doing so, the difference in the elements of alphabets $\mathcal{G}_{ii}$ could blur the general trends.

There are mainly 3 arguments in favour of discrete alphabets chosen by us:

1. Due to the payoff function - As shown in [50], for certain utilities, notably sum-rate, only Binary power control is optimal or quasi-optimal depending on the number of pairs. For sum-rate, the number of required power levels is just 2, $\{P_{min}, P_{max}\}$.
2. Feedback noise - In practical scenarios, the feedback received is noisy. One way to mitigate the noise is to quantize the feedback alphabet set. We see from Fig. 5.4 that the power levels chosen for different feedbacks g_{ii} is more uniform at higher noise levels. This points at both, less quantization points for the feedback, as well as less number of power levels to choose from.
3. Statistical estimation noise - In addition to the feedback noise, the channel statistics might not be perfectly known either. We again see that having a small alphabet set $(\mathcal{G}_i i, \mathcal{P}_i)$ indeed helps in ensuring against the possible error in estimation.

   We illustrate this in Fig. 5.7, which is interesting for many reasons. Firstly, it illustrates the robustness of our algorithm. Even in the case of $z = 1$ which corresponds to 100% error in estimation, the achieved payoffs are comparable to the case with no noise in estimation. Secondly, with more error in estimation, one requires a smaller feedback alphabet set to reach optimal performance.

   From Fig. 5.8, we see that the required card($\mathcal{P}_i$ for achieving 'good' payoffs is very less too ≈ 20. This, along with Fig. 5.7, shows that the alphabets $\mathcal{G}_{ii}, \mathcal{P}_i$ can be discrete and small without inducing much performance loss.

While in slightly more complex cases (multi-carrier, multi-user scenarios), the required alphabet sizes may be higher, we argue that due to the multitude of reasons discussed and illustrated, one expects a small enough alphabet suffices for achieving 'good' payoffs.

- ***Complexity*** : The distributed nature of the algorithm helps in reducing the complexity of the search for optimal functions. For the noiseless case, the complexity of the algorithm is $O(K.\text{card}(\mathcal{P}_i)\text{card}(\mathcal{G}_{ii})^{K^2})$. We see that it is exponential w.r.t. number of users K. While this is considerably less than the case of centralized optimization, it is still prohibitive when there are many users. To get around this issue, we firstly observe from Fig. 5.9 that the solution found using our algorithm is a combination of two strategies; thresholding and channel inversion from [1]. We see that unlike

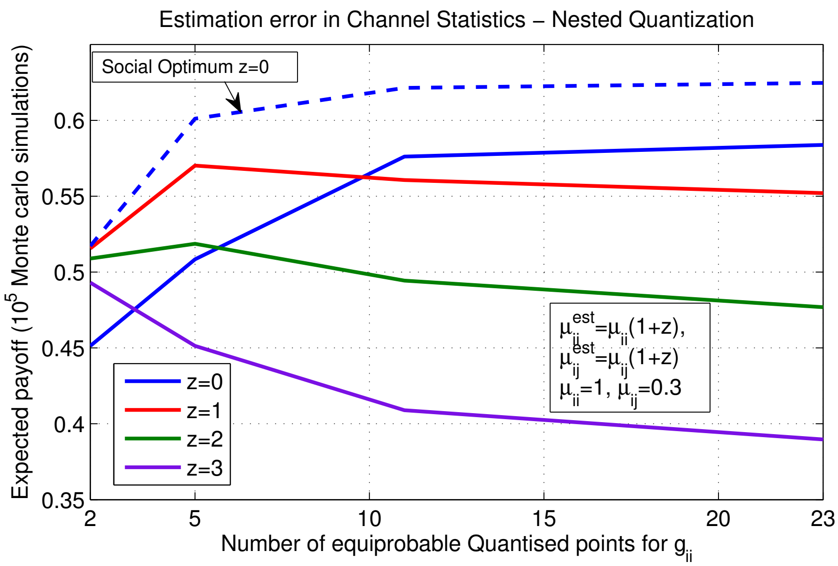


Figure 5.7: As we see from this graph, very few quantization points per channel are required to achieve payoffs comparable to a much finer quantization. When Channel statistics are estimated with an error, we notice a maximum with only few quanitzation points for different noise levels.

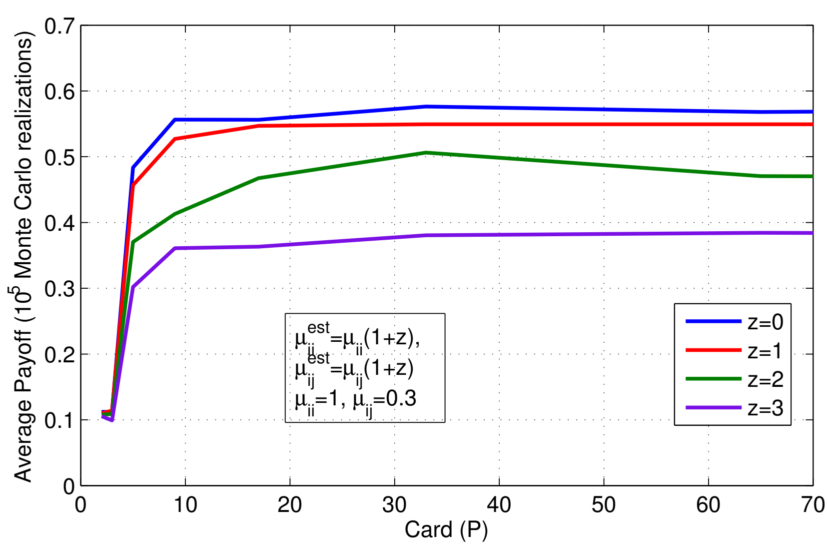


Figure 5.8: The number of Power levels required to achieve comparable payoffs to many more power levels can be seen. It can be seen that around 20 power levels are enough for providing reasonable payoffs.

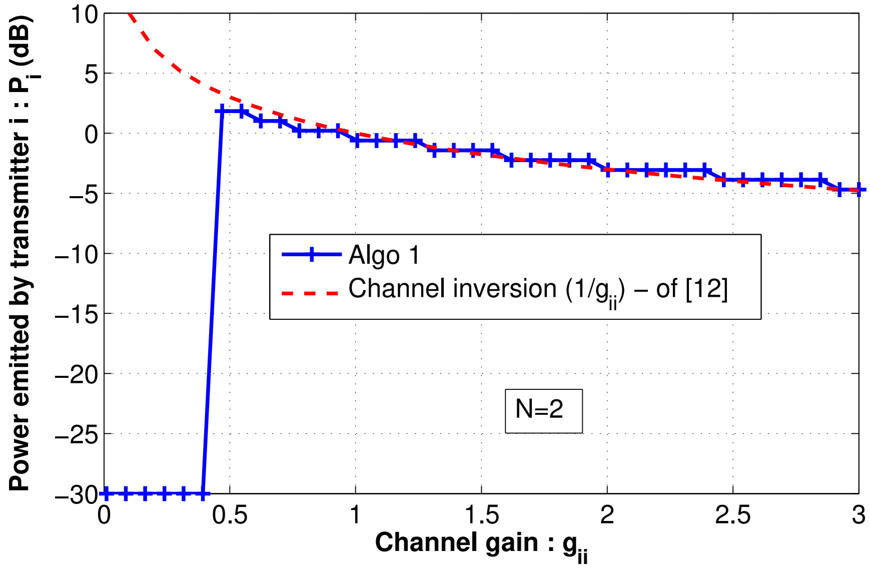


Figure 5.9: Our algorithm reveals the shape of good power control functions in the presence of interference. In contrast with related works on energy-efficiency [1], our work shows that thresholding is required to manage interference efficiently.

[1], we find a threshold below which emitting no power is more optimal, and over which the function closely follows the channel inversion suggested by [1].

Therefore, we use the intuition given by the power control policy found using the algorithm to propose the following simple continuous power control policy.

$$P_i^{\lambda_i} = \begin{cases} 0 & \text{if } g_{ii} \leq \lambda_i \\ \frac{\alpha_i}{g_{ii}} & \text{if } g_{ii} \geq \lambda_i \end{cases} \tag{5.1}$$

Using the thresholding policy, we can reduce the complexity for the search much further. In Fig. 5.10 we compare the scalability of thresholding strategy (5.1) and Nash equilibrium. Unsurprisingly, we see that our policy scales much better than Nash Equilibrium as we increase the number of TX-RX pairs. Indeed, one can show that for Nash Equilibirum, as $K \to \infty$, $\mathbb{E}w \to 0$. This proves the merit of thresholding strategy for sum-energy payoff. However, we cheated slightly here by assuming symmetric channel statistics for all TX-RX pairs which makes finding the optimal threshold very simple. In the assymetric case, exhaustive search can have prohibitve complexity beyond 3 TX-RX pairs.

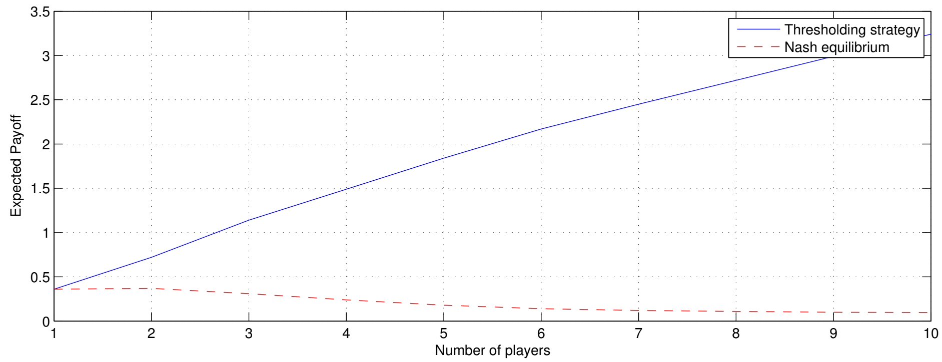


Figure 5.10: Scalability of the proposed thesholding strategy and comparison with Nash Equilibrium. We see that as number of Tx-Rx pairs increase, thresholding strategy does increasingly better, whereas Nash equilibrium payoff falls to 0

5.1.3 Concluding Remarks

The proposed framework was shown to be relevant in diverse scenarios of single band interference channel for finding optimal power control functions. Moreover, the power control functions depend only on local CSIT, thus having the merit of being implementable in a completely decentralized manner. Also, the solutions obtained take noise in the estimation of the channel gain into account. All the above features illustrate the generality of our approach in tackling problems of power control for maximizing sum-utility functions.

However, the framework can be exploited further for tackling other problems in wireless communications as well. For example, one could consider the problem of power allocation in a multi-band interference channel. Also, the auxiliary variable U was not exploited, which in general will only make the solution more optimal. Also, the thresholdng strategy could be analyzed better so as to be robust w.r.t. noisy feedback.

5.1.4 Appendix

Calculation of optimal power control function for 1 user for Energy Efficiency

Consider a general energy efficiency utility function $u_i = \frac{\Omega(\gamma_i)}{P_i}$.

At the optimal point:

$$\begin{aligned} \frac{\partial u_i}{\partial P_i} &= 0 \\ \implies \frac{\partial}{\partial P_i}\frac{\Omega(\gamma_i)}{P_i} &= 0 \\ \implies \frac{1}{P_1}\frac{-\Omega(\gamma_i)}{P_i^2} + \frac{\partial \Omega(\gamma_i)}{\partial \gamma_i}\frac{\partial \gamma_i}{\partial P_i} &= 0 \end{aligned}$$

Thus, with just one Tx-Rx pair, there exists a $\gamma_1^\star$, the solution to the equation $\gamma_i^\star \Omega'(\gamma_i^\star) - \Omega(\gamma_i^\star) = 0$, which maximises its utility.

Now, consider a specific throughput function - $\Omega(\gamma_i) = e^{\frac{-c}{\gamma_i}}$. For this case, we find that $\gamma_i^\star = c$. Thus, the optimal power control function is simply $P_i = \frac{c\sigma^2}{g_i i}$.

Nash Equilibrium Calculation - k=2

At the equilibrium point, by definition of Nash Equilibrium, Transmitter i does not gain anything by changing his transmitted power P_i, assuming the other player plays P_2. Thus

$$\begin{aligned} \frac{\partial u_i}{\partial P_i} &= 0 \\ \implies \frac{\partial}{\partial P_i}\frac{f(\gamma_i)}{P_i} &= 0 \\ \implies \frac{1}{P_1}\frac{-f(\gamma_i)}{P_i^2} + \frac{\partial f(\gamma_i)}{\partial \gamma_i}\frac{\partial \gamma_i}{\partial P_i} &= 0 \end{aligned}$$

To proceed further, we assume the utility function to be of the form $f(\gamma_i) = e^{-\frac{c}{\gamma_i}}$. The 'best response' P_1 for a given P_2 can thus be derived as follows:

$$\begin{aligned} \frac{cf(\gamma_i) - \gamma_i f(\gamma_i)}{P_i^2 \gamma_i} &= 0 \\ \implies \gamma_i(P_1^{eq}, P_2) &= c \end{aligned}$$

$$P_1^{eq} = \frac{c(\sigma^2 + P_2 g_{21})}{g_{11}} \tag{5.2}$$

Now, by symmetry, we know that $P_2^{eq} = \frac{c(\sigma^2 + P_1 g_{12})}{g_{22}}$. Substituting P_1^{eq} in the place of P_1 (because at equilibrium, both would play their equilibrium values), we get:

$$P_1^{eq} = \frac{c\sigma^2(g_{22} + cg_{21})}{g_{11}g_{22} - c^2 g_{12}g_{21}}, P_2^{eq} = \frac{c\sigma^2(g_{11} + cg_{12})}{g_{11}g_{22} - c^2 g_{12}g_{21}} \tag{5.3}$$

Clearly, depending on the realisations of g_{ij} and the constant c, one of the players might exceed the maximum power$P_{max]}$, in which case the other player will play its best response power to attain its $\gamma*$. Same logic applies if $P_i^{eq} < 0$.

5.2 Power Consumption Scheduling functions for Smart-Grids

5.2.1 Introduction

In this section, we shall divert our attention to the problem of power consumption scheduling functions for Smart-grids, which is also essentially a problem of co-ordination among the consumers to achieve certain common goals for the loads on the electricity distribution networks. With the advent of Smart Grids, it has become possible to send information via the network to the appliances and consumers and vice versa. The information thus generated and relayed could be used to schedule the consumption of electrcity (for charging or using appliances) in order to mitigate the effects of uneven consumption on the network.

An important problem for modern electrical networks is to design intelligent strategies for use of electrical appliances (Electric Vehicles (EV) being an example) which are able to exploit the knowledge they have about the non local demand or the electricity price to reach a certain objective. The objective can be to reduce the impact of electricity consumption on the distribution network or to minimize the monetary charging cost paid by the consumer. The most standard approach is to design a charging scheme which assumes a perfect knowledge of the non local demand (or price) and evaluate the performance of the corresponding algorithm by feeding the latter with a forecast or noisy version of the non local demand. Illustrative and recent examples of this approach are given e.g., by [54] [55], [56], [57]. In the quoted references, the energy need of a given user is computed by assuming perfect knowledge of the electricity price or the exogenous demand namely, the part of the demand which is not controlled by the smart consuming devices. The obtained power scheduling schemes have essentially or exactly the water-filling structure i.e., that holes in terms of price or demand are exploited in the first place. One of the drawbacks of this approach is the potential lack of robustness of the designed algorithm to imperfect forecast.

Among existing works which take uncertainty into account in the design of power scheduling scheme we find in particular [58]. Therein the authors propose a threshold-based scheduling policy which accounts for past and current prices and the statistics of future prices; each appliance consumes according to a rectangular profile and starts consuming when the price is below a time-varying threshold ("threshold policy"). In [58] the price values are assumed independent from the load level in each time slot while this assumption is relaxed in [59].

Additionally, in the latter reference the authors also consider uncertainty in the algorithm design part but the uncertainty concerns load and user energy consumption needs, and not prices as in [58]. Another example of relevant work where price uncertainty is considered in real-time demand response model is [60]; therein robust optimization is exploited. In [61], the problem is addressed by stochastic gradient based on the statistical knowledge of future prices.

To our knowledge, there is no contribution in the literature related to the present work which treats the problem of optimality of a power consumption scheduling scheme under given arbitrary imperfect observation or forecast. We use the results developed in Chap. 4 to apply to the model proposed in Sec. 5.2.2. This provides us with suboptimal but typically good power consumption scheduling schemes to be determined numerically. A thorough analysis on very practical scenarios is done in Sec. 5.2.3. In Sec. 5.2.4, the main assets of the proposed approach are summarized and several extensions to address its limitations are provided.

5.2.2 Proposed system model

We consider a set of $K \geq 1$ smart electrical appliances. Each appliance aims at scheduling its power consumption to maximize a certain payoff function, which is provided further. To this end, it exploits the available knowledge about a state which affects its payoff. In this application, this state is the *exogenous load* namely, the part of the load which is not controlled by the smart consuming devices but the proposed model and derived results can be directly used for other types of states such as the electricity price. To define the exogenous load and other key quantities such as the power consumption vectors, we need to specify the timing aspect. Time is assumed to be slotted in stages $t \in \{1, ..., T\}$. Typically, a stage may represent a day and T may represent the number of days over which the payoff is averaged. At the beginning of *stage* t, appliance k has to choose a power consumption vector $\underline{x}_k = (x_{k,1}, ..., x_{k,N})$. For example, $N = 24$ if a stage is a day and comprises 24 time-slots whose duration is one hour. The choice of consumption vector should exploit perfect observations of the past exogenous load vectors $\underline{x}_0(1), ..., \underline{x}_0(t-1)$ (note that $\underline{x}_0(t)$ is a vector of size N) and a signal which is an image or forecast of the system state and appliances actions at stage t; this signal is denoted by $\underline{s}_k \in \mathcal{S}_k$. For example, such a signal may be a forecast of the exogenous load or the total load, the total load being equal to the sum $x_0(t) + \sum_{k=1}^{K} x_k(t)$. For $k \in \{0, 1, ..., K\}$, the set in which $\underline{x}_k$ lies is assumed to be discrete and is denoted by $\mathcal{X}_k$. This choice is not only motivated by the fact that both power and time can be discrete in real systems, but also to obtain a solution robust against forecasting noise; the latter issue has been addressed recently in [2] where rectangular consumption profiles typically perform better than continuous profiles.

The computational complexity of the algorithm proposed in Sec. 4.3 depends on the cardinality of the set $\mathcal{X}_0$ (the set in which the exogenous load lies). In general, assuming $\mathcal{X}_0$ to be discrete amounts to approximating the exogenous load vectors. To obtain a good approximation, we propose to apply the principal

General model	Power consumption scheduling
Nature State - X_0	PCA analysis of $\widehat{\mathbf{R}}_L$
Decision of Agent k - X_k	Starting time to charge
$w(x_0, x_1...x_K)$	Joule Losses or Transformer Ageing
$\rho(x_0)$	Based on empirical data and PCA analysis
$\urcorner(s_k\|x_0)$	DMC with transition probability to nearest neighbour under L_2 norm

Table 5.2: Correspondence between the general framework and its application to power consumption scheduling

component analysis (PCA) [62] on exogenous load vectors. The exogenous load vector for stage t is approximated as follows:

$$\widehat{\underline{x}}_0(t) = \widehat{\underline{\mu}}_L + \sum_{i=1}^{M} a_i(t)\underline{v}_i \tag{5.4}$$

where $\widehat{\underline{\mu}}_L$ is defined by

$$\widehat{\underline{\mu}}_L = \frac{1}{L}\sum_{t\in\mathcal{L}} \underline{x}_0(t) \ , \tag{5.5}$$

$\mathcal{L}$ being a set of L samples which is available to estimate $\widehat{\underline{\mu}}_L$; typically, it may correspond to data measured during the preceding year. The vectors $\underline{v}_i$ are the eigenvectors of the following matrix

$$\widehat{\mathbf{R}}_L = \frac{1}{L}\sum_{t\in\mathcal{L}} \left[\underline{x}_0(t) - \widehat{\underline{\mu}}_L\right]\left[\underline{x}_0(t) - \widehat{\underline{\mu}}_L\right]' \tag{5.6}$$

where the notation [.]$'$ stands for transpose. One of the advantages of using such a decomposition is that for a given number of basis vectors M (the basis is then $(\underline{v}_1, \cdots, \underline{v}_M)$), the quality of approximation is maximized; more specifically, the expected distortion $\mathbb{E}\left\|\widehat{\underline{X}}_0 - \underline{X}_0\right\|^2$ is minimized. To minimize the latter quantity we will exploit the Lloyd-Max algorithm [63] in the numerical analysis; it will be applied to the vector $\underline{a} = (a_1, ..., a_M)$.

We trace the average normalised distortion $\frac{1}{L}\sum_{t\in\mathcal{L}} \|\underline{x}_0(t) - \widehat{\underline{x}}_0(t)\|^2/\|\underline{x}_0(t)\|^2$ against the number of basis vectors M that we use in our modeling of X_0 to find a 'good' value for M. $\widehat{\underline{x}}_0(t) = \widehat{\underline{\mu}}_L + \sum_{i=1}^{M} a_i\underline{v}_i$ where $\underline{a}$ is $\arg\min_{\underline{a}} \|\underline{x}_0(t) - \widehat{\underline{\mu}}_L - \sum_{i=1}^{M} a_i\underline{v}_i\|^2$ using convex optimisation algorithms. As we see from Fig. 5.11, the average distortion is more sensible to a few basis vectors $\underline{v}_i$. Thus, by accepting a distortion of 5%, we can reduce the basis space from 15 to 4 which enables us to model X_0 with fewer coefficients $\underline{a}$ and reasonable accuracy.

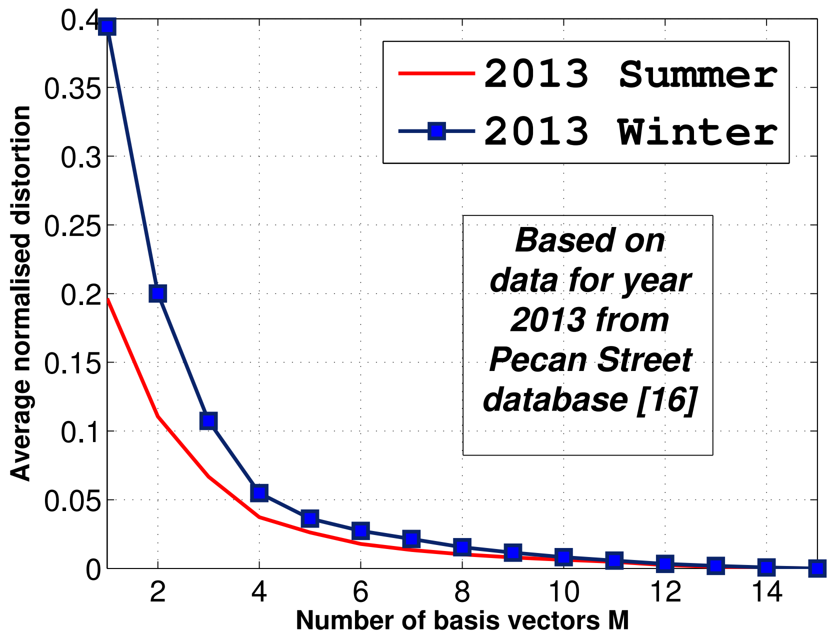


Figure 5.11: Average normalised distortion against number of basis vectors chosen. We see that 4 vectors are enough to approximate $\underline{X}_0$ for an error tolerance level of 5%.

The *stage or instantaneous payoff* function of appliance k is denoted by $u_k(\underline{x}_0, \underline{x}_1, ..., \underline{x}_K)$. In the simulations we will assume that u_k can be written as

$$u_k(\underline{x}_0, \underline{x}_1, ..., \underline{x}_K) = \sum_{n=1}^{N} u_{k,n} \Big(x_{0,1} + \sum_{k=1}^{K} x_{k,1}, \\ \cdots, x_{0,n} + \sum_{k=1}^{K} x_{k,n} \Big) \tag{5.7}$$

where n is the index for the hours of the day and $u_{k,n}$ is the consumption . Note that the cost function is not instantaneous inside a stage, but for a given day as seen from the equation 5.7. The function $u_{k,n}$ can e.g., represent the price charged to the consumer, Joule losses, battery aging, or distribution transformer aging at time-slot n. To define the average payoff of appliance k which is the function to be maximized by appliance k, the key notion of power consumption scheduling strategies needs to be defined. A *strategy* for appliance k is a sequence of mappings which is defined by:

$$\sigma_{k,t} : \begin{array}{ccc} \mathcal{X}_0^{t-1} \times \mathcal{S}_k & \to & \mathcal{X}_k \\ (\underline{x}_0(1), ..., \underline{x}_0(t-1), \underline{s}_k(t)) & \mapsto & x_k(t) \end{array} \tag{5.8}$$

The *average payoff* of appliance k is then defined by

$$U_k(\sigma_1, ..., \sigma_K) = \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T} u_k\left(X_0(t), X_1(t), ..., X_K(t)\right)\right]. \tag{5.9}$$

The expectation operator is used since the load is typically considered as a random process. Thus in full mathematical generality, the appliances decisions have also to be considered as random processes. The two main technical problems we address in this paper are the characterization of achievable average payoffs when T grows large and the determination of a practical scheme which provides good performance. Note that we assume a general structure for the knowledge $\underline{s}_k$. Indeed, we assume that $(\underline{s}_1, ..., \underline{s}_K)$ are the outputs of a general discrete memoryless channel (see e.g., [25] for more details) whose conditional probability is $\Gamma(\underline{s}_1, ..., \underline{s}_K|\underline{x}_0, \underline{x}_1, ..., \underline{x}_K)$.

Create X_0 - For good modeling of the given data to create accurate $\widehat{X_0}$, we do PCA analysis of $\widehat{\mathbf{R}}_L$ as explained in section 5.2.2. We thus obtain M 'significant' vectors. We then find the continuous vectors $\underline{a}^C(t)$ which lie in the set $\mathbb{R}^M$ and minimise the distortion $\|\widehat{\underline{X}}_0(t) - \underline{X}_0(t)\|^2$, for every day t. We discretize the continuous alphabet using vectorial Lloyd max algorithm, with the $\underline{a}^C(1), ..., \underline{a}^C(t)$ and specifying the number of points for quantization. The representatives thus generated is our alphabet $\mathcal{X}_0$.

Create noise model - Since we evaluate the average payoffs for white Gaussian noise $\underline{\widehat{x}_0} + \underline{z}$, where $Z \sim \mathcal{N}(\underline{0}, \sigma^2\mathbf{I})$, whereas our algorithm uses a DMC channel Γ, the probability of error due to the noise should be taken into account in the algorithm. To reduce the complexity, it is assumed as a first approximation that *only* the closest neighbour of $\underline{\hat{X}}_0$ in terms of the Euclidean norm $\|.\|_2$ can be gotten by error as the approximation of exogenous profiles which should have been approximated by $\underline{\hat{X}}_0$. To estimate the probability of error, 100 draws of a white gaussian noise are added to the realizations of a training set of the exogenous load over $\{1, \cdots, T\}$ and the number of times when a realization leads to the closest neighbour instead of the right approximation are counted. Then, the probability of error is estimated as the ratio of this number of errors over the number of times when a given element of $\mathcal{X}_0$ should have been chosen.

5.2.3 Numerical Analysis

Simulation setup

The simulation setup assumed here by default corresponds to a Texan district in 2013 in which the smart electrical appliances are electric vehicles. A stage is a period from 8 am a given day to 8 am the day after: there are $N = 24$ time-slots of 1 hour each. The power consumption of the smart electrical devices will be scheduled only during the "nighttime" corresponding here to the period 5 pm - 8 am the next day, i.e. $n \in \{10, \cdots, 24\}$. The whole time period is then $\{1, \cdots, T = 365\}$. Data corresponding to the exogenous profiles $(\underline{x}_0(1), \cdots, \underline{x}_0(365))$ are obtained from the Pecan Street database [64]. To get a realistic profile at the scale of a district, we sum the hourly consumption data available for Texan households and normalize the aggregated profile as explained a little further. We verified using clustering algorithms of Matlab (function kmeans) that on an average over 3 years, the real consumption of the aggregated profile in summer and winter were considerably different. Winter corresponds to $[1, 120] \cup [301, 365]$ and summer to $[121, 300]$ Thus we chose to only run the simulations for summer, with some tests for winter to see if the results of summer conform with those of winter.

In the case of EV charging, the strategy sets $\mathcal{X}_k$ are constrained by the mobility parameters and the charging need. Based on a recent French survey [65], the arrival μ_k^a, departure time μ_k^d and number of time-slots needed to charge C_k are taken to be the closest integers of realizations of Gaussian random variables $\widetilde{\mu}_k^a \sim \mathcal{N}(2, 0.75)$, $\widetilde{\mu}_k^d \sim \mathcal{N}(14.5, 0.375)$ and $\widetilde{C}_k \sim \mathcal{N}(2.99, 0.57)$. Motivated by battery aging consideration [66], it is furthermore assumed that charging profiles are rectangular: charging is done at a constant rate, here 3kW, and cannot be stopped. Note also that profiles without interruption are even required in some important scenarios encountered with home energy management [67][68]. A strategy is then define only by the time to start charging and the strategy set is reduced to $\mathcal{X}_k = \{\mu_k^a, \cdots, \mu_k^d - C_k + 1\}$, the potential time to start charging.

One important assumption in our analysis is the memorylessness of the payoff function. Regarding the instantaneous payoff function, two cases will be distinguished. The first considers a "memoryless" payoff with $u_k(\underline{x}_0, \underline{x}_1, ..., \underline{x}_K) = -\sum_{n=1}^{N}(x_{0,n} + \sum_{k=1}^{K} x_{k,n})^2$. This is convenient to express Joule losses or a price charged to the consumer depending on the total load at the scale of the district as in [54]. The second corresponds to payoffs "with memory", which means that $u_{k,n}$ depends on the whole past of the total load $(x_{0,1} + \sum_{k=1}^{K} x_{k,1}, \cdots, x_{0,n} + \sum_{k=1}^{K} x_{k,n})$ as opposed to only on the current load $x_{0,n} + \sum_{k=1}^{K} x_{k,n}$. In the context of a distribution network, such a cost can be transformer aging. To be very brief, the most influential parameter for the transformer aging is known to be the hot-spot (HS) temperature [69]. Indeed, the transformer isolation damage is directly related to the HS temperature: aging is exponentially accelerated (decelerated) when the HS temperature is above (resp. below) its nominal value [69]. This is mathematically translated as

$$\begin{aligned} & u_{k,n}(x_{0,1} + \sum_{k=1}^{K} x_{k,1}, \cdots, x_{0,n} + \sum_{k=1}^{K} x_{k,n}) \\ & = e^{0.12 \times F_t^{\mathrm{HS}}\left(x_{0,1} + \sum_{k=1}^{K} x_{k,1}, \cdots, x_{0,n} + \sum_{k=1}^{K} x_{k,n}\right) - 11}, \end{aligned} \tag{5.10}$$

which provides the instantaneous aging relatively to the nominal case. The transformer HS temperature evolution law F_t^{HS}, which has a memory, is assumed to follow the ANSI/IEEE linearized Clause 7 top-oil-rise model, which is described in [70]. To make the simulations reproducible we provide the values of the different parameters of F_t^{HS}: $\Delta_t = 0.5$ h; $T^0 = 2.5$h (thermal inertia) for all simulations concerning the transformer, $\gamma = 0.83$; $R = 5.5$; $\Delta\theta_{\mathrm{FL}}^{\mathrm{O}} = 55$ C; $\Delta\theta_{\mathrm{FL}}^{\mathrm{HS}} = 23$ C. The initial parameters for this dynamics with memory are assumed to be the same from one stage to an other given that the transformer temperature has enough time to converge to the same value at 5 pm for different values of this temperature at the beginning of the day at 8 am. in a given season. These are set to $\theta_0^{HS} = 37$ C and $x_{0,9} = 30$W (resp. $\theta_0^{HS} = 75$ C and $x_{0,9} = 87$W) in winter (resp. summer). Defining a larger number of periods than the two (winter and summer) seasons with initial parameters for each period could improve the accuracy of the estimation of the transformer lifetime but would need more data to approximate the exogenous load. With these parameters, the exogenous load is normalized such that without EV the transformer lifetime (inversely proportional to the average aging) is 40 years (standard value for nominal conditions).

The functions found using our algorithm and iterative Water Filling algorithm are applied to a noisy version of data for exogenous load (Summer 2013) and real costs are calculated assuming that the vehicles use the decision functions given by the respective algorithms. Note however, that we suppose that the vehicles already know the statistics for the future. The knowledge of future statistics was also assumed in [58]. We use some typical values for Forecasting Signal to Noise Ratio FSNR, namely FSNR = 7dB as used in [58] to be coherent with the literature. FSNR is defined as FSNR = $10\log_{10}\left(\frac{1}{\sigma_{day}^2}\frac{1}{NT}\sum_{t=1}^{T}\sum_{n=1}^{N}X_{0,n}(t)^2\right)$

Simulation Results

In our results, we distinguish 2 kinds of vehicles, *informed vehicles* and *uninformed vehicles*. By informed vehicles, we mean vehicles who receive a signal corresponding to the exogenous demand $X_0(t)$ and take their decisions accordingly, whereas the uninformed vehicles recieve no signal, and thus follow a 'plug and charge' policy, since they have no information. The total number of vehicles for the simulations was kept constant at 10. We chose to plot relative payoff loss $(u_{total}/u_{noEV}) - 1$, which measures the payoff loss when compared to the case of no EV.

In Fig. 5.12, there are 2 key messages: 1) Algo 1 is much more robust to noise, infact in Fig. 5.12, the curves for 2 different noise levels merge. IWFA however is sensible to noise, which is not surprising. Since we impose rectangular charging profiles on the vehicles, they will not react to noise since it might just be a transient optimal period to charge. 2)We see that as the number of informed players increase the difference between our algorithm and iterative Water filling algorithm (IWFA) increases.

In Fig. 5.13 we see that the general trend discussed for Fig. 5.12 is true irrespective of the year. Moreover all the 3 years show similar patterns, thus showing that the good performance at higher penetration is not by chance. We could not test for more years as that information is not available in a similar format.

We see in Table 5.3 that the ratio between the average costs of Algo 1

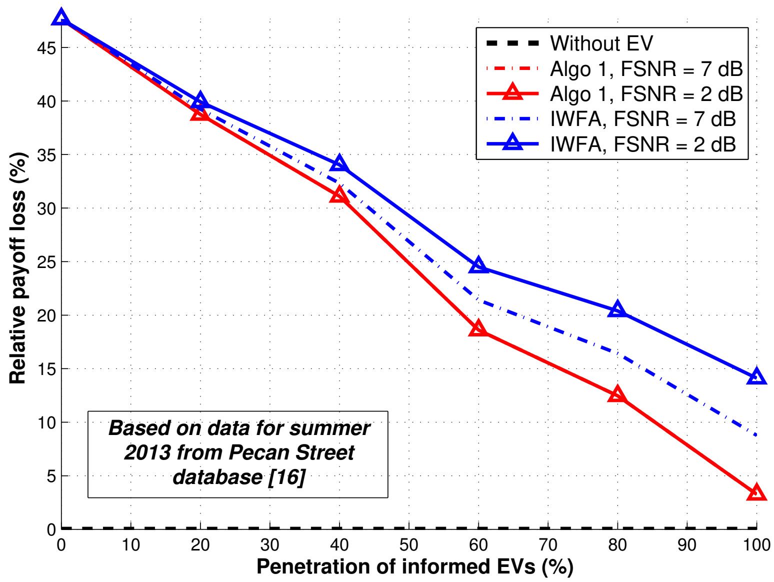


Figure 5.12: Relative payoff loss ($u_{total}/u_{noEV}) - 1$ in percent) against penetration percentage. For more discussion on the choice of this metric refer to [2]. This figure illustrates the robustness of our approach as well as increased importance at higher penetration rates.

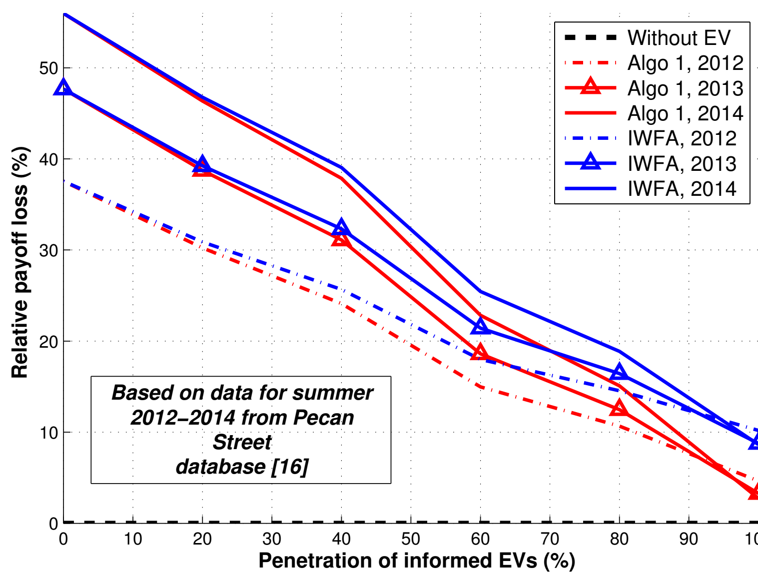


Figure 5.13: Relative payoff loss $((u_{total}/u_{noEV}) - 1$ for different years for both algorithms in comparison. We see that Algorithm 1 does considerably better every year.

Number of Vehicles	1	2	3	4	5	6	7	8
Transformer Aging	0.99	0.98	0.96	0.96	0.94	0.93	0.91	0.89
Energy Loss	0.99	0.98	0.96	0.96	0.94	0.94	0.92	0.90

Table 5.3: Ratio of u_{Algo1}/u_{IWFA} for different number of vehicles, supposing that all the vehicles are informed (in posession of a signal)

and IWFA decreases as the number of vehicles increases. This implies that as the number of vehicles increases, Algo 1 does increasingly better than IWFA. Moreover, the same pattern is observed for the cost function taken to be Energy losses. This vindicates our claim that the algorithm proposed by us is generic and could be applied to any cost function to provide 'good' strategies.

Since the cost also depends on the need $\underline{d}$ which represents the needs of the electric vehicles, we generated $\underline{d}$ using the general statistics for demand observed. We generated 3 $\underline{d}$ vectors independently with the same distribution, and found that all the results discussed above hold. Same is true for the winter periods as defined in the simulation setup. All the following tests of robustness show the general nature of our approach.

5.2.4 Concluding remarks

Numerical results show the full relevance of the proposed PCA-based model. Remarkably, an accurate approximation can be obtained by using only a few eigenvectors and applying the Lloyd-Max algorithm on the weights to be applied to these eigenvectors. The proposed framework to characterize the best performance of power consumption scheduling exploits a very recent result in information theory and also allows to derive robust scheduling functions. Simulations show that in the presence of uncertainty on the exogenous load forecast, the obtained functions outperform iterative Water-Filling based schemes. Most importantly, we provide a general framework which can be applied to various scenarios, and guarantees electric vehicle using functions achieving good performance for every scenario.

The proposed framework can be extended. While the discrete alphabet assumption seems like a good assumption to obtain robust scheduling schemes, the i.i.d assumption on the exogenous load would need to be relaxed e.g., into a milder assumption such as a Markovian process, the i.i.d. assumption being made for the performance characterization theorem. Another direction to further improve performance is to maximize the expected payoff jointly and not by using a distributed algorithm. A more accurate study with mobility data from Texas could constitute an interesting extension of this work but as a first step, we assumed that the French ones could be applied also in the context described here.

5.3 Witsenhausen Counterexample

The story will not be complete without applying theorem 4 to the Witsenhausen Counterexample (WC) which is of great interest due to the complex interplay between communication and control. Before exploring the WC, we would like to restate theorem 4 in a form more adapted to its information structure which involves agents observing each others' actions. This particularity leads to a combination of the two parts of this thesis since witsenhausen counterexample is a fundamental signalling problem involving implicit communication amongst the agents, while at the same time the information structure of the problem is causal, in the sense that no agent is aware of the future realizations of the nature state with respect to which they co-ordinate.

5.3.1 Performance limits with agents observing each other

While it may seem counterintuitive that the agents can observe each others' actions while trying to co-ordinate with only causal information at every instant t, this could happen if every t itself is composed of many stages j, $j \in \{1, ..., J\}$. For all stages j, the nature state remains constant and the payoff is decided at last stage J once all agents have made their decision. In such a scenario, an agent who needs to take a decision at stage j can, in general, observe the actions chosen by some agents during the previous stages $j \in \{1, ..., (j-1)\}$ without violating causality. By $\mathcal{Y}_i$, we denote the set of images of actions taken by all the agents that agent i can observe without violating causality before taking its decision. More formally, a more general information structure for the sequence of decision functions $f_{i,t}$ of player i for instant t could be stated as follows:

$$\begin{array}{cccll} f_{i,t} & : & \mathcal{S}_i^t \times \mathcal{U} \times \mathcal{Y}_i & \longrightarrow & \mathcal{X}_i \\ & & (s_i(1), s_i(2)...s_i(t), u(t), y_i(t)) & \longmapsto & x_i(t) \end{array} \tag{5.11}$$

where $\mathcal{S}_i^t = \mathcal{S}_i(1) \times \mathcal{S}_i(2) ... \times \mathcal{S}_i(t)$ is the discrete observation alphabet of system state x_0 upto and including the instant t, $\mathcal{X}_i$ is the decision alphabet for agent i for all instants t , with both $|\mathcal{S}_i|, |\mathcal{X}_i| < \infty$. The main difference between 4.1 and 5.11 is that agent i can also observe an image of the actions of other agents $\mathcal{Y}_i$ if permitted by the information structure. As before, $\mathcal{U}$ is the alphabet of the auxiliary variable U which could in general serve as a co-ordination key for the agents. $s_i(t)$, $u(t)$ and $y_i(t)$ are the realizations of the corresponding variables at instant t.

This leads to the generalization of Theorem 4, where we now simply add the fact that agent i can observe the actions chosen by other agents Y_i as long as the actions were chosen at a stage before agent i chooses it's decision.

Theorem 5. *Assume the random process $X_{0,t}$ to be i.i.d. following a probability distribution ρ. $X_{0,t}$ is known to agent i via its image $S_{i,t}$ which is the output of a discrete memoryless channel (DMC) obtained by marginalizing the conditional probability $\daleth_{x_0}(s_1, ..., s_N|x_0)$. Agent i also observes the actions of other agents* *within the instant t denoted by Y_i. Assume the observation Y_i to also be an output of a DMC denoted by $\Gamma_i(Y_i|X_{\mathcal{I}}) = \prod_{i' \in \mathcal{I}} \Gamma(y_{i'}|x_{i'})$, where $\mathcal{I}$ denotes the set of all agents whose actions are observable by agent i, and $X_{\mathcal{I}}$ denotes the set of actions chosen by the agents belonging to the set $\mathcal{I}$. An expected payoff $\overline{w}$ is achievable in the limit $T \to \infty$ if and only if it can be written as:*

$$\begin{aligned} \overline{w} = \sum_{\substack{x_0, x_1, \ldots x_N, \\ u, s_1, \ldots s_N}} & \rho(x_0) P_u(u) \daleth(s_1, \ldots s_N | x_0) \times \\ & \left(\prod_{i=1}^{N} P_{X_i|S_i,U,Y_i}(x_i|s_i, u, y_i) \prod_{i' \in \mathcal{I}} \Gamma_i(Y_i|X_{\mathcal{I}}) \right) \times \\ & w(x_0, x_1, \ldots x_N) \end{aligned} \tag{5.12}$$

where U is an auxiliary variable which can be optimized and $P_{X_i|S_i,U,Y_i}(x_i|s_i, u, y_i)$ is the probability that Transmitter i, chooses action x_i after observing s_i, u, y_i.

Now, we can develop an algorithm on similar lines as Algo 1, but with the added structure of the agents observing each others' actions in a perfectly causal and ordinated manner. This induces the function updates to be carried out stage wise within the instant t. Note that this does not affect the complexity of the algorithm much as all players still sequentially update their functions. After stating the algorithm for the general case of N agents and J stages per timeslot, we shall treat the special case of Witsenhausen counterexample $(N, J = 2)$.

From Theorem 5, we see that the average common payoff for the agents is characterized by the conditional probabilities of an action being chosen for each agent $P_{X_i|S_i,U,Y_i}(x_i|s_i, u, y_i)$. However, an important point to note, as has been previously explained in proposition 8, is that at optimal payoff the strategies cannot be non-deterministic. If it were the case for any agent, it could always choose to give probability 1 to the action optimising the objective to reduce the payoff. This is due to the multilinear nature of the optmization. Thus without loss in optimality, we shall restrict the search to deterministic decision functions f_i defined in 5.11. In [71] they start the optimization assuming non deterministic controllers and slowly impose determinism to obtain the optimal transport functions. While this approach is different from what we propose, the arguments used therein for imposing determinism is the same.

Writing a general algorithm for the discussed information structure in this section is pretty complex in general. Therefore, we first turn our attention to Witsenhausen counterexample which is a special case of the scenario described in Theorem 5. For this special case, an algorithm is much less convoluted and understandable. The agents are 2-controlers $(N = 2)$ co-ordinating with each realization of nature state $x_0(t)$ being i.i.d. and each instant t composed of 2 stages $(J = 2)$. In the following section, we describe the problem as well as the progress made until now in understanding the open problem.

5.3.2 Background - Witsenhausen Counterexample

The Witsenhausen Counterexample [19] provides a well studied and important coordination problem which has a hierarchy of information amongst coordinating agents. Witsenhausen showed that affine controls, even for control problems even with linear system, quadratic costs and gaussian noise - the LQG problem, may not be optimal in some cases as was previously believed. This is an artifact of the hierarchy of the information structure as explained by [72] [73].

The counterexample is a deceptively simple 2-Agent coordination problem which captures the essence of tradeoff between optimising control costs and communication costs at the same time. Reference [74] provides a very lucid description of the problem along with its importance due to the established links with computer science. However, the illustrative scenario constructed therein, of stabilizing an inverted pendulum, is too contrived in my experience of presenting the problem. For a more natural illustration, reconsider the scenario of a truck and a car coordinating to follow a path introduced in Sec. 1.2. Also, imagine that the visibility is very low, i.e. the truck can only see the immediate path ahead. The truck has two main objectives, follow the path, as well as signal the path to the car behind given that the car too has a blurry observation of the movement of truck. The communication-control trade-off comes in because manipulating a big and heavy truck is costly.

The trade-off can be better understood by considering extreme cases. Imagine the truck is really heavy. In that case, the truck will simply follow the path and let the car behind guess the path ahead based on its blurry observation of the truck's movements. On the other hand, if the truck is ultra-light and easily manipulable, it will not be too costly for it to move more than required taking into account how blurry the vision of the car behind is. Thus, based on how heavy the truck is, and how blurry the vision of the car is, they have to choose the right strategy for coordination. It turns out that finding optimal strategies for coordination problems with such information structure is fiendishly difficult, if not impossible.

The problem can be formally stated in the following manner (See Fig. 5.14).

- Uncontrolled signals - x_0, z where x_0 represents the initial system state and z represents the noise in the observation for the second agent. System state $x_0 \sim \mathcal{N}(0, \sigma_0^2)$ whereas the noise is normalized $z \sim \mathcal{N}(0, 1)$.
- Decision functions - First agent observes the system state perfectly and chooses its action $x_1 = u(x_0)$. The second agent has a noisy observation $y = x_1 + z$ of the first agent's action and chooses its action $x_2 = v(y)$.
- Common Objective - Choose control functions u, v to minimize the expected cost : $\min\limits_{u,v} w = \mathbb{E}[k(x_1 - x_0)^2 + (x_2 - x_1)^2]$, $k \geq 0$

In the case of classical information structure, i.e. Agent 2 is privy to the information available at Agent 1, the optimal strategy is trivial. Agent 1 would, in that case, choose to take no action, resulting in $x_1 = x_0$ and zero first stage cost. Since Agent 2 knows x_0, and by extension x_1, it would simply choose $x_2 = x_0$, which would result in the second stage cost to be zero too. Thus these solutions are trivially linear. This is sometimes also referred to as centralized optimization in the literature. A natural conjecture was that linear controllers

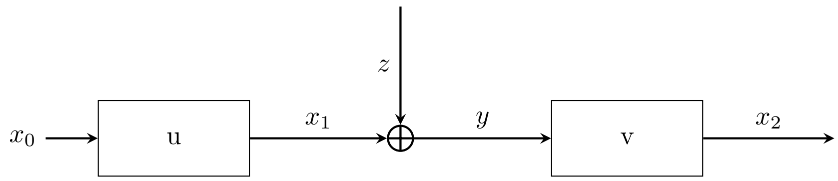


Figure 5.14: Schematic diagram of Witsenhausen counterexample

are optimal even when Agent 2 does not know the observation at Agent 1. Witsenhausen debunked this conjecture by firstly finding the optimal linear controllers and their corresponding costs in this scenario, and then showing that for certain values of parameter$k$, a simple non-linear controller can achieve lower costs for certain values of the parameters k, σ_0^2. The non linear control strategy proposed by Witsenhausen was a 1-bit quantizer of the state x_0 into either σ_0 or $-\sigma_0$. Eventhough Witsenhausen in [19] proved the existence of an optimal solution, it has, until now, not been found.

Instead, the difficulty in finding optimal control functions was further highlighted by [75] which showed that the discrete version of the problem is NP-complete. NP-complete problems is a class of problems for which no efficient solutions have yet been found. Due to their proven equivalence, if an efficient solution is found for any NP-complete problem, it would immediately resolve the other NP-complete problems.

Nonetheless various heuristic based approaches have been suggested over the years . The convergence of the costs of these sub-optimal approaches for the benchmark case $k = 0.2$, $\sigma_0 = 5$ (Fig. 5.15) leads us to believe that they are very close to optimal controllers. However, most approaches give no guarantee of their optimality. Most works use the following observations made in the original Witsenhausen's article [19] to simplify the problem.

- For an optimal strategy u, $\mathbb{E}[u(x_0)] = 0$ and $\mathbb{E}[u(x_0)^2] \leq 4\sigma_0^2$.
- Given a decision function u respecting the above criterion, the optimal decision for second agent $v(\cdot)$ is simply the MMSE:

$$v_u^\star = \mathbb{E}[u(x_0)|y] = \frac{\mathbb{E}_{x_0}[u(x_0)\phi(y - u(x_0)]}{\mathbb{E}_{x_0}[\phi(y - u(x_0)]} \tag{5.13}$$

  where $\phi(\cdot)$ is the normalized gaussian density function.

The optimal expected cost, taking into account these observations thus reduces to a function of just the decision of Agent 1.

$$w^\star(u) := w^\star(u, v_u^\star) = k\mathbb{E}[(x_0 - u(x_0))^2] + 1 - I(D_u) \tag{5.14}$$

where $I(D_u)$ is the Fischer Information of the variable y

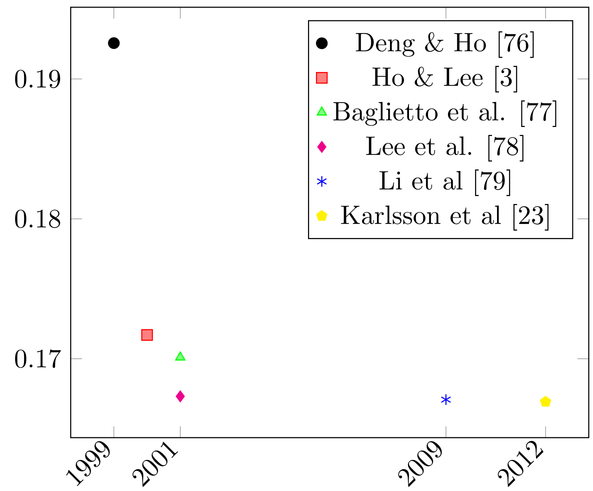


Figure 5.15: Evolution of minimum payoffs achieved by different approaches over the last 20 years. We see that the minimum has not reduced much since 2001, leading us to suspect that, atleast for the benchmark case, we might be close to optimal.

$$I(D_u) = \int (\frac{d}{dy} D_u(y))^2 \frac{dy}{D_u(y)} \tag{5.15}$$

$$D_u(y) = \int \phi(y - u(x_0))\phi(x_0; 0, \sigma_0^2) dx \tag{5.16}$$

where $\phi(x; 0, \sigma_0^2)$ is the gaussian density function with mean 0 and variance σ_0^2. Most heuristics thus search only for the optimal function $u(\cdot)$, as the optimal $v_u^\star(\cdot)$ is fixed given optimal $u(\cdot)$.

Some of the important early works analyze the reasons for the difficulty of the optimization problem given the non-classical information structure. In [22], the authors consider a slightly modified version of the Gaussian test channel (GTC), which has the same information structure as Witsenhausen counterexample, but admits optimal linear solutions. More precisely, if we consider the following payoff with the same information structure, the optimal controlers are linear:

$$w = \mathbb{E}[k(x_1 - x_0)^2 + (x_2 - x_0)^2] \tag{5.17}$$

The authors then consider a more general decentralized control problem with the GTC and WC being special cases. They show that linear strategies are optimal in the case when there is no product-term of the decision variables in the cost function. They argue that the presence of such a product term is the reason for non-linear strategies outperforming linear strategies. To prove their result, they consider the problem as a communication problem where first agent is the encoder trying to transmit a gaussian source over a gaussian channel. An important step in their approach is the use of information-theoretic data-processing inequality to convexify the problem, making it tractable. [29] argue that one could use a more general optimization approach for solving problems with non-classical information structure. They suggest transforming the problem into another equivalent optimization problem with linear objective but necessarily non-convex constraints. Thereafter, one could use the technique of convex relaxation to 'convexify' the constraints and find a solution for the new optimization problem. If then, one can find a solution which is also feasible for the original non-convex constraints, then one has found the solution. They show that the data processing inequality used in [22] does precisely this.

Various studies have shed more light on the properties of the optimal control functions. In particular, from the vantage of optimal transport theory, [80] shows that the optimal function $u(\cdot)$ is monotonically increasing and has a real analytic left-inverse. This disproved the conjecture in [81] that optimal solution is a multi-step piecewise constant function.

Notwithstanding the intractability of the problem, attempts have been to find best possible solutions from various viewpoints. [76] supposes the decision function $u(\cdot)$ to be a multi-step function, and uses ordinal optimization to narrow down the search for optimal number of steps and break-points. Ordinal optimization is a numerical method which uses quick and rough estimates using monte-carlo simulations over the parameter space to narrow down regions of interest. They assume that $u(\cdot)$ is an odd function and consider only the positive domain of the problem. They find that 2-step functions (4 steps over the whole domain) have a higher probability of 'good' payoffs and go on to finetune the search assuming 2-step functions. [3] takes this approach a step further using 2 new ideas. Firstly, they use integration with the appropriate step size to evaluate the cost of a given function, thus speeding up the search allowing them to analyze more cases of step-functions. Also, they analyze each constant section of the step function to find more 'sub-steps' thus reducing the cost further. The best solution proposed by them is a 3.5 step solution. [77] use approximating networks to find similar performances as [3] , albeit their method is much more general. They also apply their method to [22] as a verification test of their method.

More recently, [79] modeled the problem in terms of potential games considering the intervals where the step function remains constant as players trying to optimize a common utility. They then applied the procedure of joint fictitious state play (JFSP) to converge to best known solutions then. However, due to the large size of 'players', it required many more iterations to converge, although the cost dropped considerably after the first 100 iterations. A similar approach, albeit not directly inspired, was taken in [23] where they consider each controller as a player co-ordinating with the other trying to maximize a common utility. Their approach bears a close resemblance to ours, although we arrive at similar conclusions from different roots. We shall elaborate the connections in greater detail after having developed our approach in the next section. We develop a general framework of finding good coordination between N-agents, given a particular order of decision making and information available to the agents. We apply this framework to the special case of Witsenhausen counterexample.

5.3.3 Algorithm for searching optimal Witsenhausen control functions

We can now use theorem 5 to devise an algorithm for the Witsenhausen counterexample on similar lines as Chapter 4. As it would probably be clear by now, witsenhausen counterexample consists of 2 stages, i.e. $J = 2$. Also, the first controller perfectly observes the nature state $S_1 = X_0$, whereas the second controller only observes a noisy image of the action chosen by the first controller, i.e. $Y_2 = X_1 + Z$, where $Z \sim \mathbb{N}(0, \sigma_Z^2)$. While it was difficult to write a general algorithm like we did in Chapter 4, in the specific case of witsenhausen counterexample, this task is considerably simpler.

Algorithm 2: Proposed decentralized Algorithm for finding decision functions for Witsenhausen Counterexample

inputs : $\mathcal{X}_0, f_i^{\text{init}}, \forall i \in \{1, 2\}$,
output: f_1^*, f_2^*

Initialization: $f_1^0(s_1) = s_1, f_2^0(y_2) = (y_2), \quad iter = 0, \quad iter_{max} = 100$

while $\exists i : f_i^{\text{iter}-1} - f_i^{\text{iter}} \geq \epsilon \ \ AND \ \ iter \leq iter_{max} \qquad OR \qquad \text{iter} = 0$
do

  iter = iter + 1;

  foreach $x_0 \in \mathcal{X}_0$ **do**

    $f_1^{\text{iter}}(x_0) = \arg\min_{x_1 \in \mathcal{X}_0} \sum_{y_2} P(y_2|x_1) w(x_0, x_1, f_2^{iter}(y_2))$;

  end
  foreach $y_2 \in \mathcal{X}_0$ **do**

    $f_2^{iter}(y_2) = \mathbb{E}_{X_0}(x_1|y_2)$;

  end
end

The algorithm proposed is essentially the same as [23]. However, we implement the same idea in a slightly different way. The major differences are

- **Expectation over** X_0 - In [23], they take the expectation by generating 10^6 monte carlo samples of X_0. We argue that, as shown in [24], one could just create a finite model of X_0 by uniform quantization over a finite range. We argue that doing this, we can achieve good enough performances for a much smaller complexity.

- **Range of** X_0 - In [23], the range of X_0 is taken to be $[-25, 25]$, without any justification. We ran our algorithm for different ranges and found that the range has a role to play in the performance of the algorithm. [3] also performs ordinal optimization to ascertain the range as well as stepsize of X_0 although they use a different procedure for finding the functions.

- **Parameter Relaxation** - In [23], due to potentially bad initialization points, they adopt the technique of parameter relaxation. Essentially, the parameter k, i.e. the weight of the cost associated at first stage, is initially taken to be high and slowly decreased to the desired value 0.04, in approximately 10 steps. We argue that such an arbitrary step is not required, and one can simply initialize with identity linear controllers i.e. $u(x_0) = x_0$. We see that this converges to equally good solutions as [23].

- **Assumption of symmetry** - In [23], they assume that u is an odd function, i.e. $u(x_0) = -u(-x_0)$. Since the solutions obtained are not necessarily optimal, this restriction could, in general, lead to worse solutions. We impose no such restriction, and indeed our solutions do not verify this assumption close to $x_0 = 0$.

5.3.4 Numerical analysis

As mentioned in the preceding discussion, the implementation of Algo 4.3 differs in our case. To make a just comparison, we define certain quantities pertaining to the range and stepsize of the alphabets $\mathcal{X}_i$. The range for the nature alphabet $\mathcal{X}_0$ is taken between $[-M, M]$, with the stepsize δ_{X_0}. The range for $\mathcal{X}_1, \mathcal{X}_2$ is taken to be the same. However, their stepsize is decreased after Algo 4.3 converges and the algo is ran again. This is done by preserving the decision functions found beforehand and merely extending their domains to include the mid-points between the earlier domain points.

Firstly, we would like to point out the variability of average payoff on the range M as well as the stepsize δ_{X_0}. As seen in Fig. 5.16, the average payoff obtained are highly variable based on the chosen range and stepsize. At first glance, the result might seem paradoxical, but considering a bigger range for X_0 is not necessarily helpful as it might increase the chances of extra cost due to noise in observation of the second agent. Thus, there is a sort of Braess Paradox in choosing the range.

While we did not carry out a thorough analysis to find good range and stepsize, by preliminary simulations, we identified a case ($M = 16.47$,$\delta_{X_0} = 0.27$) which gives a good payoff 0.1745 when $\mathcal{X}_0 = \mathcal{X}_1 = \mathcal{X}_2$. Then, we increase the alphabets $\mathcal{X}_1, \mathcal{X}_2$ by reducing the corresponding stepsize by half. Let us denote the number of total points of $|\mathcal{X}_i| = \frac{2M}{\delta_{X_i}} = L$.

With the notations in order, we can compare our results with that of [23] for different alphabet sizes L. As we see from 5.17, for lower L, we provide better results, simply by searching for 'good' range and stepsize to begin with. While it is not clear if this trend holds true at higher L, we assert that it is still

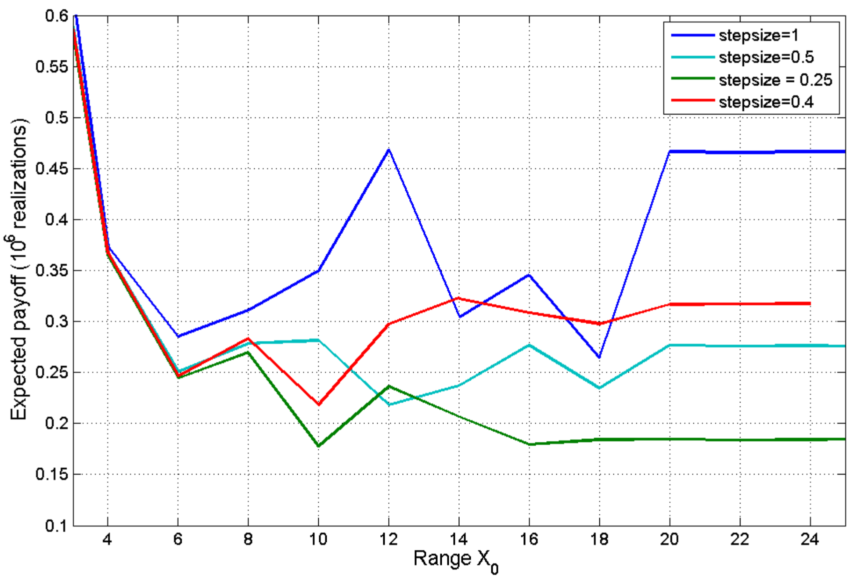


Figure 5.16: Vagaries of average payoff w.r.t. range as well stepsize of the alphabet $\mathcal{X}_0$. As we see, there is no fixed pattern. If one considers a bigger range for X_0, conventional wisdom says that we must do better, but it is not necessarily the case. We note also that for all stepsizes, above a certain range, the payoff remains constant . We conclude that one needs to employ techniques akin to ordinal optimization [3] to find optimal range and stepsize.

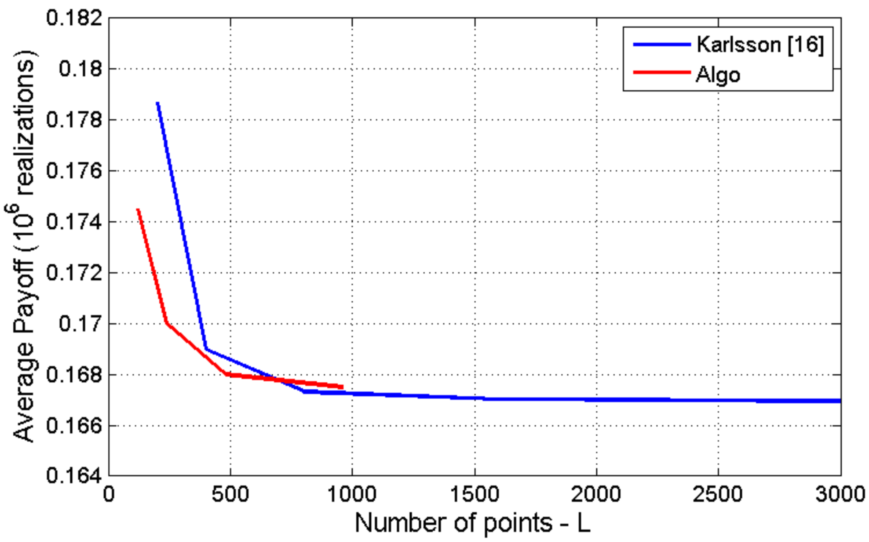


Figure 5.17: We see that if one chooses the right range and stepsize to begin with, one could obtain much better performances for low number of quantization points L. Unfortunately, we have not been able to go further in the simulations to be able to compare the two implementations at higher L. Nonetheless, searching for a 'good' initial range and stepsize can help reach better performances for lower complexity

worthwhile to perform an initial search for range and stepsize, especially if one is looking for solutions which have low complexity.

5.3.5 Concluding Remarks

While the algorithm developed in this section had already been found independently before, we provide a rationale behind why it is a good algorithm. However, due to the sub-optimality of the approach, we could not attack the more important and open problem of finding the optimal decision functions as well as lowest achievable costs. Nonetheless, the approach outlined in this section is in all probability $\epsilon-$optimal due to the result in [24]. Also, slight differences in the application of the algorithm provides hope of beating the lowest known solution till date if one manages to find the right range to search over, and given enough time to run extensive simulations. Given the really marginal improvements in the cost for the benchmark case however, it is not clear if such an effort is worthwhile, even though it would easily merit a publication in a reputed conference.

Chapter 6

Conclusions

This thesis has merely scratched the surface of the possibilities opened up by the Information theoretical tools for evaluating performances of coordination schemes. One hopes that it has nonetheless demonstrated the potential of those tools by exploring applications in various, otherwise unrelated, domains. Eventhough we tried being clever in modeling the problems to exploit the tools developed herein, in all applications further advancements were restricted due to complexity issues.

Even otherwise, there is a need to develop the theory further than the current status as coordination problems do not always assume the information structures discussed in this thesis. For example, the non-causal information structure considered in the first part currently only applies to $2-$Agent team problems, which is evidently too restrictive. There is a need to consider a more general information structure. One possibility could be to consider one agent having non-causal information about the future realizations of the nature state, while others observe its action to coordinate with it. This problem is much more complicated than the $2-$agent case, as the 'informed' agent, in addition to performing implicit communication, also has to choose the importance it attaches to communication with each follower. Also, the follower needs to guess if the action chosen by the informed agent was meant for him. All this to say that while the extension to this case is non trivial, if one wants to apply the developed theory to realistic coordination problems with non-causal information structure, it is necessary to go beyond 2-Agents.

Another possibility, probably a simpler case to analyze, would be to consider a chain of agents observing each other, with the leading agent having non-casual information. Even then, the performance limit of such a system is not easily obtainable from the current theory, as one cannot simply write the information constraints for each implicit communication in the chain. Moreover, even though this might be a simpler case to consider, for application purposes, the assumed information structure is too restrictive. In this thesis, we willingly did not develop the applications for the non-causal information structure in great detail as theoretically, there is still work to be done.

For the second part pertaining to causal information structure, the challenges lie more on the application side. The developed theory is pretty general and easily adaptable to various applications as evidenced from chapter 5. However, there seems to be a trade-off between optimality and complexity which is difficult to know a priori and is highly dependent on the application in question. This requires careful analysis and modelling of the problems, and good knowledge of the state of the art to be able to choose the right mix of optimality and complexity to provide solutions better than the existing ones.

Theoretically too, there is some progress to be made, especially concerning the information structure discussed for the Witsenhausen Counterexample, where each timeslot could be divided into various stages, and agents could observe each others', causality permitting. This should be a slightly easier endeavour, although any practical application based on this theory will necessarily suffer from complexity issues, as evidenced from the treatment of the relatively simple (complexity wise) $2-$agent Witsenhausen counterexample.

Most importantly, there are many new domains to be explored which could use the framework developed to provide practical coordination schemes. An application, which I find interesting, and which undoubtedly is going to be useful in the future is that of coordination for controlling the flow of traffic. A simple toy problem could be to consider traffic signals co-operating with each other with the common objective of maximizing traffic flow. One can imagine with the development in sensor technology that it would be possible for signals to locally measure the traffic in each direction, and with very little back-haul, relay the information to neighbouring signals. The signals could then make informed and smart decisions about how many vehicles to let go in which direction. This is but one such example of hopefully many more to come.

Bibliography

[1] D. Goodman and N. Mandayam, "Power control for wireless data," *Personal Communications, IEEE*, vol. 7, no. 2, pp. 48–54, Apr 2000.

[2] O. Beaude, S. Lasaulce, M. Hennebel, and I. Mohand-Kaci, "Reducing the Impact of EV Charging Operations on the Distribution Network," *Smart Grid, IEEE Trans. on*, 2015 (revised), Available: https://drive.google.com/file/d/0BwhXJd7d5TqoeGo3dG1NOHljQjA/-view?usp=sharing).

[3] Y.-C. Ho and J. T. Lee, "Granular optimization: An approach to function optimization," in *Decision and Control, 2000. Proceedings of the 39th IEEE Conference on*, vol. 1. IEEE, 2000, pp. 103–111.

[4] J. F. Nash, "Equilibrium points in n-person games," *Proceedings of the National Academy of Sciences*, vol. 36, no. 1, pp. 48–49, 1950. [Online]. Available: http://www.pnas.org/content/36/1/48.short

[5] J. Marschak, "Elements for a theory of teams," *Management science*, vol. 1, no. 2, pp. 127–137, 1955.

[6] R. Radner, "Team decision problems," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 857–881, 1962.

[7] B. Larrousse, S. Lasaulce, and M. Wigger, "Coordination in state-dependent distributed networks: The two-agent case," in *Information Theory (ISIT), 7th IEEE International Symposium on*, 2015.

[8] ——, "Coordinating partially-informed agents over state-dependent networks," in *Information Theory Workshop (ITW), 2015 IEEE*, April 2015, pp. 1–5.

[9] B. Larrousse, S. Lasaulce, and M. Bloch, "Coordination in distributed networks via coded actions with application to power control," *arXiv preprint arXiv:1501.03685*, 2015.

[10] P. Grover, "Actions can speak more clearly than words," 2010.

[11] E. Chan, P. Gilhead, P. Jelinek, P. Krejci, and T. Robinson, "Cooperative control of sartre automated platoon vehicles," in *Proceedings of the 19th ITS World Congress*, 2012, pp. 22–26.

[12] O. Gossner, P. Hernandez, and A. Neyman, "Optimal use of communication resources," *Econometrica*, vol. 74, no. 6, pp. 1603–1636, 2006.

[13] I. Khayutin, "Communication with unobservable constraints," in *2013 IEEE International Symposium on Information Theory (ISIT'2013)*, Jul. 2008.

[14] 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.

[15] 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, 2008.

[16] 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*, November 2010.

[17] 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, August 2011, special Issue on Game Theory for Wireless Networks.

[18] S. Lasaulce and H. Tembine, *Game Theory and Learning for Wireless Networks : Fundamentals and Applications*, Elsevier, Ed. Academic Press, August 2011.

[19] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, pp. 131–147, 1968.

[20] B. Larrousse and S. Lasaulce, "Coded power control: Performance analysis," in *2013 IEEE International Symposium on Information Theory (ISIT'2013)*, Jul. 2013.

[21] B. Larrousse, S. Lasaulce, and M. Bloch, "Coordination in distributed networks via coded actions with application to power control," *submitted to IEEE Transaction on Information Theory*, 2014.

[22] R. Bansal and T. Başar, "Stochastic teams with nonclassical information revisited: When is an affine law optimal?" *IEEE Transactions on Automatic Control*, vol. 32, no. 6, pp. 554–559, June 1987.

[23] J. Karlsson, A. Gattami, T. Oechtering, and M. Skoglund, "Iterative source-channel coding approach to witsenhausen's counterexample," in *American Control Conference (ACC), 2011*, June 2011, pp. 5348–5353.

[24] N. Saldi, S. Yüksel, and T. Linder, "Finite model approximations and asymptotic optimality of quantized policies in decentralized stochastic control," *arXiv preprint arXiv:1511.04657*, 2015.

[25] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* Wiley-Interscience, 2006.

[26] S. P. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge university press, 2004.

[27] R. Mochaourab and E. Jorswieck, "Resource allocation in protected and shared bands: uniqueness and efficiency of nash equilibria," in *Proceedings of the Fourth International ICST Conference on Performance Evaluation Methodologies and Tools.* ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2009, p. 68.

[28] B. Larrousse and S. Lasaulce, "More about optimal use of communication resources," in *International Workshop on Stochastic Methods in Game Theory, Erice, Italy*, Sep. 2013.

[29] A. A. Kulkarni and T. P. Coleman, "An optimizer's approach to stochastic control problems with nonclassical information structures," in *51st IEEE Conference on Decision and Control (CDC'2012)*, Dec. 2012.

[30] M. H. Costa, "Writing on dirty paper (corresp.)," *Information Theory, IEEE Transactions on*, vol. 29, no. 3, pp. 439–441, 1983.

[31] T. M. Cover and J. A. Thomas, *Elements of information theory.* John Wiley & Sons, 2012.

[32] A. El Gamal and Y.-H. Kim, *Network information theory.* Cambridge University Press, 2011.

[33] L. A. Steen and J. A. . Seebach, *Counterexamples in Topology*, 1970.

[34] R. Gray, *Probability, Random Processes, and Ergodic Properties.* SpringerVerlag, 2009.

[35] A. Wyner, "A definition of conditionnal mutual information for arbitrary ensemble," *Inform. and Control*, pp. 51–59, 1978.

[36] 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.

[37] B. Larrousse, A. Agrawal, and S. Lasaulce, "Implicit coordination in 2-agent team problems. application to distributed power allocation," in *IEEE 12th Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, May 2014.

[38] N. A. Ahmed and D. Gokhale, "Entropy expressions and their estimators for multivariate distributions," *IEEE Transactions on Information Theory*, vol. 35, no. 3, pp. 688–692, 1989.

[39] A. G. Nahapetyan, "Bilinear programming bilinear programming," in *Encyclopedia of Optimization*. Springer, 2009, pp. 279–282.

[40] S. Lasaulce and H. Tembine, *Game theory and learning for wireless networks: fundamentals and applications*. Academic Press, 2011.

[41] R. Gohary and T. Willink, "Robust IWFA for open-spectrum communications," *Signal Processing, IEEE Transactions on*, vol. 57, no. 12, pp. 4964–4970, Dec 2009.

[42] A. Anandkumar, A. Anandkumar, S. Lambotharan, and J. Chambers, "Robust rate maximization game under bounded channel uncertainty," *Vehicular Technology, IEEE Transactions on*, vol. 60, no. 9, pp. 4471–4486, Nov 2011.

[43] D. Monderer and L. S. Shapley, "Potential games," *Games and economic behavior*, vol. 14, no. 1, pp. 124–143, 1996.

[44] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *Vehicular Technology, IEEE Transactions on*, vol. 41, no. 1, pp. 57–62, 1992.

[45] R. Yates, "A framework for uplink power control in cellular radio systems," *Selected Areas in Communications, IEEE Journal on*, vol. 13, no. 7, pp. 1341–1347, Sep 1995.

[46] S. Stanczak, M. Wiczanowski, and H. Boche, *Fundamentals of resource allocation in wireless networks: theory and algorithms*. Springer Science & Business Media, 2009, vol. 3.

[47] A. Agrawal, S. Lasaulce, O. Beaude, and R. Visoz, "A framework for optimal decentralized power control with partial csi," 2015.

[48] A. Gjendemsjø, G. E. Øien, and D. Gesbert, "Binary power control for multi-cell capacity maximization," in *Signal Processing Advances in Wireless Communications, 2007. SPAWC 2007. IEEE 8th Workshop on*. IEEE, 2007, pp. 1–5.

[49] D. G. Paul de Kerret., Samson Lasaulce and U. Salim, "Best-response team power control for the interference channel with local CSI," in *IEEE International Conference on Communications (ICC)*, June 2015.

[50] A. Gjendemsj, D. Gesbert, G. Oien, and S. Kiani, "Binary power control for sum rate maximization over multiple interfering links," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 8, pp. 3164–3173, August 2008.

[51] G. Scutari, D. Palomar, and S. Barbarossa, "Asynchronous iterative water-filling for gaussian frequency-selective interference channels," *Information Theory, IEEE Transactions on*, vol. 54, no. 7, pp. 2868–2878, July 2008.

[52] F. Meshkati, M. Chiang, H. Poor, and S. Schwartz, "A game-theoretic approach to energy-efficient power control in multicarrier CDMA systems," *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 6, pp. 1115–1129, June 2006.

[53] E. V. Belmega and S. Lasaulce, "Energy-efficient precoding for multiple-antenna terminals," *Signal Processing, IEEE Transactions on*, vol. 59, no. 1, pp. 329–340, 2011.

[54] A.-H. Mohsenian-Rad, V. W. S. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous Demand Side Management Based on Game-Theoretic Energy Consumption Scheduling for the Future Smart Grid," *Smart Grid, IEEE Trans. on*, vol. 1, no. 3, pp. 320–331, 2010.

[55] M. Shinwari, A. Youssef, and W. Hamouda, "A water-filling based scheduling algorithm for the smart grid," *Smart Grid, IEEE Trans. on*, vol. 3, no. 2, pp. 710–719, 2012.

[56] L. Gan, U. Topcu, and S. H. Low, "Optimal decentralized protocol for electric vehicle charging," *Power Systems, IEEE Trans. on*, vol. 28, no. 2, pp. 940–951, 2013.

[57] O. Beaude, C. Wan, and S. Lasaulce, "Composite charging games in networks of electric vehicles," in *Network Games, Control and Optimization (NetGCooP), 7th International Conference on*, 2014.

[58] T. T. Kim and H. V. Poor, "Scheduling power consumption with price uncertainty," *Smart Grid, IEEE Trans. on*, vol. 2, no. 3, pp. 519–527, 2011.

[59] P. Samadi, H. Mohsenian-Rad, V. W. S. Wong, and R. Schober, "Tackling the Load Uncertainty Challenges for Energy Consumption Scheduling in Smart Grid," *Smart Grid, IEEE Trans. on*, vol. 4, no. 2, pp. 1007–1016, 2013.

[60] A. J. Conejo, J. M. Morales, and L. Baringo, "Real-time demand response model," *Smart Grid, IEEE Trans. on*, vol. 1, no. 3, pp. 236–242, 2010.

[61] R. Deng, Z. Yang, J. Chen, and M.-Y. Chow, "Load Scheduling With Price Uncertainty and Temporally-Coupled Constraints in Smart Grids," *Power Systems, IEEE Trans. on*, vol. 29, no. 6, pp. 2823–2834, 2014.

[62] S. Mallat, *A wavelet tour of signal processing.* Academic press, 1999.

[63] S. Lloyd, "Least squares quantization in pcm," *Information Theory, IEEE Trans. on*, vol. 28, no. 2, pp. 129–137, 1982.

[64] (2013) Public data basis. Pecan Street energy data (Texas). [Online]. Available: https://dataport.pecanstreet.org/.

[65] (2008) Public data basis. French ministry of Ecology, Sustainable Development and Energy. [Online]. Available: http://www.statistiques.developpement-durable.gouv.fr/transports/trv/deplacement-mobilite/mobilite-reguliere-locale.html.

[66] L. Gan, U. Topcu, and S. H. Low, "Stochastic distributed protocol for electric vehicle charging with discrete charging rate," in *IEEE Power and Energy Society General Meeting*, vol. 27, no. 3, 2012, pp. 1–8.

[67] T. Bapat, N. Sengupta, S. K. Ghai, V. Arya, Y. B. Shrinivasan, and D. Seetharam, "User-sensitive scheduling of home appliances," in *Proceedings of the 2nd ACM SIGCOMM Workshop on Green Networking*. ACM, 2011, pp. 43–48.

[68] P. Chavali, P. Yang, and A. Nehorai, "A distributed algorithm of appliance scheduling for home energy management system," *Smart Grid, IEEE Trans. on*, vol. 5, no. 1, pp. 282–290, 2014.

[69] IEEE, *Guide for loading mineral-oil-immersed transformers*. IEEE Std. C57.91-1995, 1995. [Online]. Available: http://books.google.fr/books?id=PvFEPQAACAAJ

[70] L. J. Rivera and D. J. Tylavsky, "Acceptability of four transformer top-oil thermal models: Pt. 1: defining metrics," *Power Delivery, IEEE Trans. on*, vol. 23, no. 2, pp. 860–865, 2008.

[71] Y. Wu and S. Verdu, "Witsenhausen's counterexample: A view from optimal transport theory," in *50th IEEE Conference on Decision and Control (CDC'2012)*, Dec. 2011.

[72] Y. C. Ho and K.-h. Chu, "Team decision theory and information structures in optimal control problems–part i," *Automatic Control, IEEE Transactions on*, vol. 17, no. 1, pp. 15–22, 1972.

[73] K.-C. Chu, "Team decision theory and information structures in optimal control problems–part ii," *Automatic Control, IEEE Transactions on*, vol. 17, no. 1, pp. 22–28, 1972.

[74] A. Sahai and P. Grover, "Demystifying the witsenhausen counterexample [ask the experts]," *Control Systems, IEEE*, vol. 30, no. 6, pp. 20–24, Dec 2010.

[75] C. H. Papadimitriou and J. Tsitsiklis, "Intractable problems in control theory," *SIAM journal on control and optimization*, vol. 24, no. 4, pp. 639–654, 1986.

[76] Y.-C. Ho *et al.*, "An ordinal optimization approach to optimal control problems," *Automatica*, vol. 35, no. 2, pp. 331–338, 1999.

[77] M. Baglietto, T. Parisini, and R. Zoppoli, "Numerical solutions to the witsenhausen counterexample by approximating networks," *Automatic Control, IEEE Transactions on*, vol. 46, no. 9, pp. 1471–1477, Sep 2001.

[78] J. T. Lee, E. Lau, and Y.-C. Ho, "The witsenhausen counterexample: A hierarchical search approach for nonconvex optimization problems," *Automatic Control, IEEE Transactions on*, vol. 46, no. 3, pp. 382–397, 2001.

[79] N. Li, J. R. Marden, and J. S. Shamma, "Learning approaches to the witsenhausen counterexample from a view of potential games," in *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on.* IEEE, 2009, pp. 157–162.

[80] Y. Wu and S. Verdú, "Witsenhausen's counterexample: A view from optimal transport theory." in *CDC-ECE*, 2011, pp. 5732–5737.

[81] P. Grover, A. Sahai, and S. Y. Park, "The finite-dimensional witsenhausen counterexample," in *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, 2009. WiOPT 2009. 7th International Symposium on.* IEEE, 2009, pp. 1–10.

ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre : Développement et application des bornes issues de la théorie de l'information à certains types de problèmes de coordination.

Mots clés : Theorie de l'information, Réseaux Sans Fil, Smart Grids, Contre-exemple de Witsenhausen

Résumé : Avec la montée de la connectivité entre les appareils (internet des objets), nouvelles possibilités de coordination entre les différentes entités ont ouvert. En même temps, des résultats récents, issus de la théorie de l'information, ont fourni des limites pour la performance que tout système de coordination pourrait atteindre sous certaines structures d'information. Dans cette thèse, nous développons ces résultats théoriques dans le but de les rendre plus facilement applicable aux problèmes pratiques. À cet égard, la contribution de cette thèse est double: 1) En outre développer les résultats théoriques pour fournir un aperçu de la structure des solutions au problème d'optimisation posés dans les travaux anterieurs , ainsi que la généralisation des résultats . 2) Développer des algorithmes qui exploitent le cadre théorique fourni par les travaux antérieurs pour concevoir des mécanismes de coordination pratiques, décentralisées et robustes. La généralité de l'approche se prête à diverses applications, dont les éléments suivants ont été traités : optimisation de puissance dans les réseaux sans fil, planification de la consommation d'énergie dans les applications de réseau intelligent, ainsi que Witsenhausen contre-exemple, un problème important issu de la théorie du contrôle. Diverses possibilités sont encore à venir pour exploiter le cadre et les outils développés ici. En effet, ils pourraient être utiles même dans des domaines qui ne sont pas abordés dans cette thèse, mais qui nécessitent une coordination entre les agents avec des informations différentes à la disposition de chacun.

Title : Development and Application of Information Theoretical Bounds to Certain Class of Coordination Problems

Keywords : Information Theory, Wireless Networks, Smart Grids, Witsenhausen Counterexample

Abstract : With the rise in connectivity between appliances (Internet of Things), new avenues for coordination between various entities have opened up. At the same time, recent information theoretical results have provided bounds for the performance that any coordination scheme could achieve under certain information structures. In this thesis, we further develop those information theoretical results with the aim of making them applicable more easily to practical problems. In this regard, the contribution of this thesis is twofold: 1) Further developing the aforementioned information theoretical results to provide insights into the structure of the solutions to optimization problem posed in them, as well as generalizing some results. 2) Developing algorithms which exploit the theoretical framework provided by Information theory to devise practical, decentralized and robust coordination schemes. The generality of the approach lends itself to various applications, of which the following were treated: power optimization in wireless networks, power consumption scheduling in smart grid applications, as well as Witsenhausen counterexample, an important toy problem in control theory. Various opportunities still lie ahead to exploit the framework and tools developed herein. Indeed, they could be useful even in domains which have not been explored in this thesis but which require coordination between agents with different information available to each.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France