



HAL
open science

Distributed Learning versus Information Sharing in the Smart Grid: A Learning Game Approach

Hélène Le Cadre, Jean-Sébastien Bedo

► **To cite this version:**

Hélène Le Cadre, Jean-Sébastien Bedo. Distributed Learning versus Information Sharing in the Smart Grid: A Learning Game Approach. 2012. hal-00740893v1

HAL Id: hal-00740893

<https://hal.science/hal-00740893v1>

Preprint submitted on 11 Oct 2012 (v1), last revised 15 Mar 2016 (v9)

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.

Collaborative Learning is Better

Hélène Le Cadre* Jean-Sébastien Bedo†

July 7, 2012

Abstract

In this article, we focus on the identification of emerging economic organizations while agents are learning hidden individual sequences modeling renewable energy production and microgrid instantaneous needs in a decentralized hierarchical network. The network is made of 3 categories of agents: producers, providers and end users belonging to microgrids. In this uncertain context, providers are penalized in case where they cannot satisfy the entire demand of the associated microgrid. Identically, producers are penalized in case where they cannot deliver the quantity of energy booked by the providers. Service providers need to make efficient forecasts about the hidden individual sequences to optimize their decisions concerning the quantities of energy to book and the prices of the energy. We prove that there exists prices that provide to the producers a guarantee to avoid penalties. Additionally, under external regret minimization, collaborative learning through a grand coalition where the providers share their information and align their forecasts, enables them to minimize their average loss. As an illustration, we compare the convergence rates of the collaborative learning strategy with rates resulting from selfish learning based on external and internal regret minimization in a 2 producers, 3 providers network. The results confirm the theory: collaboration is better for the providers.

Keywords: Distributed Learning; Regret; Algorithmic Game Theory; Coalition

1 Introduction

In this article, we describe and test distributed learning algorithms in the context of a hierarchical highly connected network of agents. The main objective of this article is to identify the economic organizations that might emerge as a result of the learning process i.e., *collaborative learning* in a grand coalition or in sub-coalitions. We have chosen to place the model in the context of the smart grid although the results that we derive are quite generic and can be applied to any hierarchical network with a distributed access to the scarce resource and no capacity storage. Such framework is quite common in the revenue management literature [15] and can be applied to many other

*CEA, LIST, 91191 Gif-sur-Yvette CEDEX, FRANCE *E-mail*: helene.lecadre@gmail.com

†Orange Research center, 6, Place d'Alleray, 75015 Paris, FRANCE *E-mail*: jean-sebastien.bedo@orange.com

industries like to model the interconnection relations between autonomous systems or content providers and access providers in the telecommunications, to better understand the relations between suppliers and retailers in the retail industry, etc.

Going back to the smart grid, we give a broad definition of it. Initially, smart grids are networks enabling a decentralized storage of the energy and involving bidirectional energy flows which are controlled by a complex, global and secured communication network. The network is said to be *smart* because it is capable of integrating efficiently each agent's action (producers, providers and end users) in order to guarantee a sustainable and secure supply at lower cost. The main goals of the smart grid are:

- To facilitate the integration of the decentralized and intermittent production such as renewable energies.
- To get active the consumer within the electrical system.
- To provide the consumer with a set of signals allowing its effective steering consumption.
- To significantly reduce the environmental footprint of the electrical system.

Traditional electrical networks rely on nuclear based energies. The main difficulty in such a case is to adapt the production level so as to meet the uncertain demand level. We have built a first model in [4] where two learning strategies based on tit for tat and fictitious play are used to adapt the production level to meet the demand level. In this context, the distributed learning algorithms can be utilized as distributed control algorithms that provide strong asymptotic guarantees on the emergent global behavior. As output, we were able to evaluate each agent's financial gain and his bias in the prediction process. Because of their structure, smart grids offer a substantial opportunity for the integration of renewable energies. Renewable energies like wind, solar, geothermal, biomass, small hydroelectric facilities, etc., are highly unpredictable since they depend on uncontrollable exterior factors like wind, level of sunshine, etc. They are cleaner but their production being more difficult to forecast, they are far more difficult to integrate into the electrical network. However, the governments' objective is to include at least 33% of renewable energies in 2020.

The challenge associated to the understanding of such a global complex system is double. First, it requires to develop efficient learning algorithms to predict both the end users' instantaneous needs, which can be highly erratic due to their new active role in the grid, and decentralized renewable energy production. Second, it is necessary to develop an economic model capturing the intrinsic motivations and complex hierarchical relations between the involved agents. To perform this latter point, we resort to use game theory which enables us to derive analytically the agents' optimal decisions in prices and traded quantities of energy.

In practice, the electrical network based on the smart grid model is composed of a multitude of microgrids. Microgrids are modern, small-scale versions of the centralized electrical system. They can be either sellers in case where they have a surplus

of power to transfer, or buyers in case where they need to buy additional power to meet their demand. Saad et al. propose a distributed microgrid coalition formation algorithm enabling an improvement of 31% of the average power losses relative to the non-cooperative case, in [13]. However, their result relies on strong assumptions. Indeed, they make the hypothesis that the consumers' demand is random, which is quite common in practical smart grid networks, since it depends on unpredictable factors such as consumption level, consumption behavior, etc. [7]. Additionally, they make the simplifying assumption that the power surplus which is defined as the difference between the total power and the demand, is distributed according to a known density function.

Actually, the justification of the fitting of a specific parametric density function requires the game designer to learn at least its parameters. In the statistical learning literature, there are three major learning approaches, each one of them corresponding to a particular abstract learning task. These are: supervised learning, unsupervised learning and reinforcement learning [2], [16]. Tasks that fall within the paradigm of reinforcement learning are control and online optimization problems, games and other sequential decision making tasks. Learning based on regret minimization as described in [2], belongs to this category. Additionally, we observe in [5] that the performances resulting from learning based on regret minimization tested on real data bases, are clearly superior to the ones obtained using supervised learning approaches. As a result, these points have convinced us to use a learning approach based on regret minimization. The difficulty is then to extend the already existing method to a distributed learning framework and to clearly identify the emerging global behaviors.

The existing literature on distributed learning primarily focuses on distributed learning algorithms that are suitable for implementation in large scale engineering systems [6], [11], [16]. The results mainly concentrate on a specific class of games, called games of potential [14]. This class of games is of particular interest since they have inherent properties that can provide guarantees on the convergence and stability of the system. However, there exists some limitations to this framework. The most striking one is that it is frequently impossible to represent the interaction framework of a given system as a potential game [8].

The learning game studied in this paper belongs to the category of *repeated uncoupled games* since one player cannot predict the forecasts and so actions of the other agents at a given time period. To take his decision i.e., optimal prices and traded quantities of energy, each agent is aware of the history of forecasts of all the agents and of his utility. Recent work has shown that for finite games with generic payoffs there exists completely uncoupled learning rules i.e., rules where the agents observe only their own prediction history and their utility, that lead to Nash equilibria that are Pareto optimal [11]. Marden et al. exhibit a different class of learning procedures that lead to Pareto optimal vector of actions that do not necessarily coincide with Nash equilibria [8]. The capacity for a learning algorithm to converge toward a Pareto equilibrium is particularly interesting especially for practical applications. Indeed, in many problems the Nash equilibrium is inefficient i.e., its reaching leads to outcomes that are far less profitable for the agents than those resulting from a Pareto equilibrium. However,

reaching the Pareto frontier is difficult since it requires the agents to bargain and share information using complex communication mechanisms [9]. Under conditions stating that it is not possible to divide the interacting agents into two distinct subsets that do not mutually interact with one another, the game dynamics induce a Markov process over the finite state space which is defined as the set of the triples containing the chosen action, the resulting utility and an additional binary parameter called the agent's mood. Marden et al. focus on characterizing the support of the limiting stationary distribution i.e., the stable states. In particular, they prove that any stable state maximizes the social welfare under an initial assumption on agent interdependence [8]. The main difference with our model is that in their paper, the agents have the possibility to control their production level. They take as an example the wind turbines which can adapt their power to maximize the whole wind farm's social welfare defined as the sum of the total power produced by each turbine. On the contrary, in our game setting, the production of the renewable energies cannot be controlled since it relies on exogenous events. Therefore, it requires to introduce decentralized learning approaches based on online optimization [1] and to study the resulting economic organization using a game theoretic approach.

The article is organized as follows. In Section 2, we introduce the economic basis of our model, the agents and their utilities. Then the complete information Stackelberg game is solved in Section 3 proceeding by backward induction. We derive analytically the optimal prices and traded quantities of energy for the agents. Partial information is introduced in Section 4 where the interacting agents learn in a distributed fashion hidden individual sequences. Using the theoretical results obtained in the previous sections, we explain how to simulate the smart grid behavior in Section 5 and an illustration of the previously derived theoretical results is provided for a 2 producers, 3 providers network. Finally, we conclude in Section 6.

Notations

In the table below, we have listed the main notations used throughout the article.

s_i	Service provider i
\mathcal{M}_i	Microgrid attached to service provider s_i
e_k	Energy producer k
$\tilde{p}_k(t)$	e_k 's price for one unity of energy
$q_{ik}(t)$	Quantity of energy booked by s_i to e_k
$p_i(t)$	s_i 's price for one unit of energy
$\nu_i^e(t)$	Microgrid \mathcal{M}_i energy needs
$a_i(t)$	Microgrid \mathcal{M}_i decision variable
$\nu_k^e(t)$	e_k 's production in renewable energies
$\tilde{\gamma}_i$	Coefficient of the penalty imposed to e_k by s_i
γ_i	Coefficient of the penalty imposed by microgrid \mathcal{M}_i to s_i
$\pi_i(t)$	s_i 's utility
$\alpha_{ki}(t)$	Proportion of its resource allocated by e_k to s_i
$C_k(t)$	$\sum_{j=1, \dots, n} \tilde{\gamma}_j q_{jk}(t)$
$\tilde{\pi}_k(t)$	e_k 's utility

δ	$\sum_{j=1,\dots,n} \frac{1}{\gamma_j \bar{\gamma}_j}$
$(x)_+$	$\max\{x; 0\}$
$L(i)$	$1 - \frac{n-1}{\delta \gamma_i \bar{\gamma}_i}$
$f_i(\nu_i^s, t)$	s_i 's forecast about the energy needs of \mathcal{M}_i
$f_i(\nu_k^e, t)$	s_i 's forecast about e_k 's production
$f_i(t)$	Set of forecasts made by s_i at t
$f(t)$	Set of forecasts made by all the providers at t
$f_{-i}(y, t)$	Set of forecasts of all the providers except s_i who predicts y at t
$\nu(t)$	Set of the energy needs of all the microgrids and of the productions of all the producers at t
\mathcal{X}_e	Space of values for the produced energy
\mathcal{X}_s	Space of values for the consumption of the microgrids
$l_i(f(t), \nu(t))$	s_i 's loss function
$l(f(t), \nu(t))$	Meta-player's loss
T	Horizon of the repeated game
$\mathcal{R}_i(T)$	s_i 's external regret
$\mathcal{R}(T)$	Meta-player's external regret
$d_i(X)$	Randomized strategy for forecast X
$l_i^{(1)}(f_i(t), \nu(t))$	s_i 's partial loss
$\tilde{l}_g(f(t), \nu(t))$	Sum of the providers' partial losses
\tilde{V}_g	Value of the game with partial losses as utilities
$l_g(f(t), \nu(t))$	Sum of the providers' losses
$H_X(x, t)$	Payoff associated to forecast X evaluated in x at t
F_s	Set of all the predictors for each service provider
F_m	Set of all the predictors for the meta-player
\mathcal{V}_t	Sum of the variances associated with payoff $H_X(X_t, t)$
$f_C(\nu_i^s)$	Grand coalition forecast about the energy needs of \mathcal{M}_i
$f_C(\nu_k^e)$	Grand coalition forecast about e_k 's production

2 The model

The number of agents interacting in the smart grid is large. In this article, we model the smart grid as a 3 layer hierarchical network which evolution depends on the interactions between the agents composing each layer and also on the ability of the agents to cope with unexpected random events resulting in renewable energy production and microgrid energy need variations. In this section, we detail the 3 main categories of agents and the economic relations between them. Then under complete information, the repeated game which captures the interactions between the 3 layers of the hierarchical network is described. Finally, we go deeper in the agents individual characterization to identify the decision variables, the information available to each agent and their utilities.

2.1 Description of the agents

We model the smart grid through 3 types of agents: the end users, the service providers and the energy producers.

The end users have some energy needs and fulfill them either by buying energy to a service provider or by finding alternatives (solar panels, more efficient appliances, etc.). Each service provider has the possibility to buy energy to all energy producers, transport and resell it to the end users. Each energy producer produces and sells energy to all service providers.

We assume that each end user contracts with only one service provider and does not churn from one service provider to another during all the period of our study. This assumption holds well if we consider local or regional utility companies for example. In this sense, the set made of provider and attached consumers can be seen as an individual microgrid, as defined in [12] and recalled in the Introduction. We denote by s_i , with i varying between 1 and n , the service providers and by \mathcal{M}_i the corresponding group of end users. The energy producers are denoted e_k with k varying between 1 and K . The energy producers can be associated with nuclear plants, photovoltaic park managers, wind farm administrators, etc. The produced energy is supposed renewable. Non-renewable energy producers like nuclear plants might be integrated into the grid. It requires to use distributed control rules as the ones described in [4], [8]. In this article, we consider exclusively renewable energy producers.

The economic relations between the agents in the grid is pictured in Figure 1. The symbol \$ is used to represent the directed monetary transfers between the involved agents.

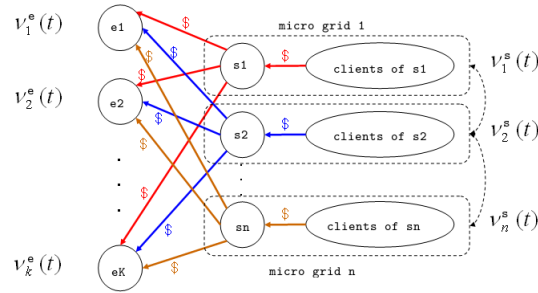


Figure 1: Economic relations between the agents in the grid.

2.2 Repeated game setting

We model the interplay between all the agents through a repeated game. At each time period t of the game¹:

- (1) The energy producers e_k communicate their prices $\tilde{p}_k(t) > 0$ for one energy unit (i.e., Wh) to the service providers
- (2) The service providers s_i place energy quantity orders to energy producers: the quantity ordered by s_i to e_k is denoted by $q_{ik}(t)$
- (3) The service providers s_i communicate their prices $p_i(t) > 0$ for one energy unit to their end users

The end users \mathcal{M}_i need $\nu_i^s(t)$ energy units for the period (which could depend on weather, cooking, etc.)

- (4) The end users decide to find alternative sources of energies for $a_i(t)$ energy units

The end users buy the rest of their needs $(\nu_i^s(t) - a_i(t))_+$ to service provider s_i

Each energy producer e_k produces $\nu_k^e(t)$ energy units

The energy producers distribute their production to service providers: e_k delivers $\alpha_{ki}(t)\nu_k^e(t)$ energy units to service provider s_i

★ The energy producers incur penalties if they did not fulfill the energy orders placed by the service providers ★

★ The service providers incur penalties if they did not fulfill the energy needs of their end users ★

The penalties are proportional to the difference between the initial energy order and the final energy delivery. More precisely, e_k incurs the penalty $\tilde{\gamma}_i \left(q_{ik}(t) - \alpha_{ki}(t)\nu_k^e(t) \right)_+$ with $\tilde{\gamma}_i > 0$, from service provider s_i , and s_i incurs the penalty $\gamma_i \left(\nu_i^s(t) - a_i(t) - \sum_k \alpha_{ki}(t)\nu_k^e(t) \right)_+$ with $\gamma_i > 0$, from end users \mathcal{M}_i . The penalties are donated to the unbiased regulator who is supposed to control the overall energy distribution system. Various mechanisms of transfer can then be implemented by the unbiased regulator. The penalties can be distributed to shareholders of the energy producers or used for social purposes.

2.3 Optimization program for each agent

In the rest of the section, we will describe for each time period t the decision variables, the available information and finally the utilities for each category of agents.

¹The cyan color is used to capture unexpected random events whereas the red one is associated to the penalty rules.

2.3.1 End users

The only decision variable for the end users is the quantity of energy that they decide to get from alternative sources: $a_i(t)$. We assume that the end users have no lever to influence their random energy needs: $\nu_i^s(t)$.

The information available to end users is the price of an energy unit from their service provider and their energy needs $\nu_i^s(t)$ for the time period.

We assume that finding alternative energy sources rather than buying it to the service provider has some cost for the end users. More precisely, finding $a_i(t)$ energy units through alternatives costs them $\frac{a_i(t)^2}{2}$ per time period. As a result, the total cost of energy for the end users \mathcal{M}_i is:

$$p_i(t) \left(\nu_i^s(t) - a_i(t) \right) + \frac{a_i(t)^2}{2} \quad (1)$$

End users \mathcal{M}_i choose $a_i(t)$ in order to minimize their total cost of energy depending on the energy price.

2.3.2 Service providers

The decision variables for each service provider s_i are the energy unit price $p_i(t)$ and the energy orders $q_{ik}(t)$ for each energy producer e_k .

The only information available to service provider s_i when he makes his decision are the energy unit prices $\tilde{p}_k(t)$ of all the energy producers. He has to forecast the energy needs of his customers and the energy production of all the energy producers.

Following our description of the interplay between the agents, the utility for service provider s_i at time period t is:

$$\begin{aligned} \pi_i(t) = & p_i(t) \left(\nu_i^s(t) - a_i(t) \right) - \sum_{k=1, \dots, K} q_{ik}(t) \tilde{p}_k(t) - \gamma_i \left(\nu_i^s(t) - a_i(t) \right) \\ & - \sum_{k=1, \dots, K} \alpha_{ki}(t) \nu_k^e(t) \end{aligned} \quad (2)$$

Service provider s_i chooses his energy unit price and his energy orders toward energy producers so that $\pi_i(t)$ is maximized.

2.3.3 Energy producers

The only decision variable for each energy producer e_k is the energy unit price $\tilde{p}_k(t)$ that he proposes to the service providers. We assume that the energy producer cannot influence directly the energy he produces at each time period. This assumption holds well if, for example, we look at a wind turbine farm without any investment in an additional wind turbine during the study period. The variation of the wind intensity will make vary the energy produced without any lever for the energy producer.

When energy producer e_k makes his decision, he has no information because he is the first player to play in the time period as described in the game setting in Section 2.2. He has to forecast the energy quantity that he will produce and the energy orders of all the service providers.

To define the sharing coefficients $\alpha_{ki}(t)$, we consider a weighted proportional allocation of resource that allows producers to discriminate energy allocation by providers. This framework is a generalization of the well-known proportional allocation [13] to *weighted energy orders* with penalty coefficients as weights. Such a resource sharing mechanism has already been introduced by Nguyen and Vojnović, in [10]. This means that between two providers booking the same quantity, the one having the highest penalty coefficient will receive the largest part of the producer's available energy. Indeed, the producer wants to minimize his overall penalty and therefore allocates larger parts of his production to providers who appear to him as more threatening than the others. More precisely, in the rest of the article, we will assume that:

$$\alpha_{ki}(t) = \frac{\tilde{\gamma}_i q_{ik}(t)}{C_k(t)} \quad (3)$$

where $C_k(t) = \sum_{j=1, \dots, n} \tilde{\gamma}_j q_{jk}(t)$.

Then the utility of energy producer e_k at time period t equals:

$$\tilde{\pi}_k(t) = \tilde{p}_k(t) \sum_{i=1, \dots, n} q_{ik}(t) - \sum_{i=1, \dots, n} \tilde{\gamma}_i q_{ik}(t) \left(1 - \frac{\tilde{\gamma}_i}{C_k(t)} \nu_k^e(t)\right)_+ \quad (4)$$

Energy producer e_k chooses his energy unit price so that $\tilde{\pi}_k(t)$ is maximized.

3 Complete information game resolution

The game setting described in Section 2.2 implies that in the relation producers-providers, producers appear as leaders whereas providers are followers. Identically, in the relation providers-consumers, providers appear as leaders whereas consumers are mere followers. Under such a setting, the game is called a Stackelberg game and as usual, it should be solved using backward induction [9].

Additionally, we make the assumption that each energy producer receives at least one energy order from a service provider guaranteeing that the Stackelberg game admits non trivial solutions.

3.1 Optimization of the end users' decision

To minimize their total cost of energy defined by Equation (1), end users \mathcal{M}_i have to choose $a_i(t)$ so that the differentiate of the total cost of energy equals 0 which means:

$$a_i(t) = p_i(t) \quad (5)$$

3.2 Optimization of the service providers' decisions

To find his optimal price and energy orders, service provider s_i has to replace $a_i(t)$ by its optimal value in $\pi_i(t)$ defined in Equation (2), and to differentiate the result in $p_i(t)$ and in $q_{ik}(t)$. This differentiation raises two cases.

3.2.1 Case 1: the energy production fulfills the energy demand of the end users

It is the case when:

$$\nu_i^s(t) - p_i(t) \leq \sum_{k=1, \dots, K} \alpha_{ki}(t) \nu_k^e(t) \quad (6)$$

Then differentiating the service provider's utility in $q_{ik}(t)$ leads to:

$$\frac{\partial \pi_i(t)}{\partial q_{ik}(t)} = -\tilde{p}_k(t)$$

which means that s_i will try to minimize all his energy orders to maximize his utility. As a result, s_i will tend to break the inequality defining Case 1 in Inequality (6) because $\alpha_{ki}(t)$ will tend toward zero. As a result the optimal decision for s_i will always fall in Case 2 described below or on the frontier between Case 1 and Case 2. The frontier between these two cases is defined by the equation:

$$\nu_i^s(t) - p_i(t) = \sum_{k=1, \dots, K} \alpha_{ki}(t) \nu_k^e(t) \quad (7)$$

3.2.2 Case 2: the energy production does not fulfill the energy demand of the end users

It is the case when $\nu_i^s(t) - p_i(t) \geq \sum_{k=1, \dots, K} \alpha_{ki}(t) \nu_k^e(t)$. Then differentiating s_i 's utility gives us:

$$\begin{aligned} \frac{\partial \pi_i(t)}{\partial p_i(t)} &= \nu_i^s(t) + \gamma_i - 2p_i(t) \\ \frac{\partial \pi_i(t)}{\partial q_{ik}(t)} &= -\tilde{p}_k(t) + \gamma_i \nu_k^e(t) \frac{\partial \alpha_{ki}(t)}{\partial q_{ik}(t)} \end{aligned} \quad (8)$$

By using the definition of $\alpha_{ki}(t)$ given in Equation (3), we obtain:

$$\frac{\partial \alpha_{ki}(t)}{\partial q_{ik}(t)} = \tilde{\gamma}_i \frac{C_k(t) - \tilde{\gamma}_i q_{ik}(t)}{C_k(t)^2}$$

Then going back to System of equations (8), we conclude that the differentiates equal 0 when:

$$p_i(t) = \frac{\nu_i^s(t) + \gamma_i}{2} \quad (9)$$

$$\tilde{p}_k(t) C_k(t)^2 = \gamma_i \nu_k^e(t) \tilde{\gamma}_i (C_k(t) - \tilde{\gamma}_i q_{ik}(t)) \quad (10)$$

On one side, we obtain directly the price for which the differentiate of $\pi_i(t)$ equals 0 through Equation (9). On the other side, Equation (10) can be rewritten as follows:

$$\tilde{\gamma}_i q_{ik}(t) = C_k(t) - \frac{\tilde{p}_k(t) C_k(t)^2}{\nu_k^e(t) \gamma_i \tilde{\gamma}_i} \quad (11)$$

If s_i anticipates that the other service providers will make the same optimization program, replicating Equation (11) for the n service providers and summing them all, results in the following equality:

$$C_k(t) = n C_k(t) - \frac{\tilde{p}_k(t) C_k(t)^2}{\nu_k^e(t)} \sum_{j=1, \dots, n} \frac{1}{\gamma_j \tilde{\gamma}_j}$$

by definition of $C_k(t)$.

Then as $C_k(t)$ is not zero because each producer e_k receives at least one order of energy otherwise he would be out of the game, by dividing the previous equation by $C_k(t)$ and reordering we obtain:

$$C_k(t) = \frac{\nu_k^e(t) n - 1}{\tilde{p}_k(t) \delta}$$

where $\delta = \sum_{j=1, \dots, n} \frac{1}{\gamma_j \tilde{\gamma}_j}$. By replacing $C_k(t)$ in Equation (11), we obtain the energy orders for which the differentiates of $\pi_i(t)$ equals 0:

$$q_{ik}(t) = \frac{\nu_k^e(t) n - 1}{\tilde{p}_k(t) \delta \tilde{\gamma}_i} L(i) \quad (12)$$

where we have introduced the notation $L(i) = 1 - \frac{n-1}{\delta \gamma_i \tilde{\gamma}_i}$ to simplify future calculations.

Presently, we have to check that the price and energy orders for which the differentiates of $\pi_i(t)$ equal 0 satisfy the conditions of Case 2.

First, it is easy to check that the price is positive through Equation (9). However, the energy orders defined in Equation (12) are non-negative if, and only if, $1 \geq \frac{n-1}{\delta \gamma_i \tilde{\gamma}_i}$ which is equivalent to:

$$\gamma_i \tilde{\gamma}_i \geq \frac{n-1}{\delta} \quad (13)$$

This inequality means that the penalties related to s_i are close to the penalties related to the other service providers. Indeed, if all penalties are equal to γ , then $\delta = \frac{n}{\gamma^2}$ and Inequality (13) is true for all service providers. On the contrary, if all penalties are equal to γ except for s_1 which has a penalty of $\frac{\gamma}{n-1}$, then $\delta = \frac{(n-1)n}{\gamma^2}$ and Inequality (13) becomes $n \geq (n-1)^2$ which is false as soon as $n > 2$.

Second, by replacing the energy orders defined by Equation (12) in Equation (3), we obtain $\alpha_{ki}(t) = \frac{L(i)}{\sum_{j=1, \dots, n} L(j)} = L(i)$ meaning that the total energy delivered to the customers of s_i is $\sum_{k=1, \dots, K} \alpha_{ki}(t) \nu_k^e(t) = L(i) \sum_{k=1, \dots, K} \nu_k^e(t)$. As a result,

the price and energy orders for which the derivatives of $\pi_i(t)$ equal 0 verify the inequality defining Case 2 if, and only if:

$$\nu_i^s(t) \geq \gamma_i + 2L(i) \sum_{k=1, \dots, K} \nu_k^e(t) \quad (14)$$

This inequality states that the total production of energy by energy producers should not be too large compared to the energy needs of customers.

If Inequalities (13) and (14) are true, the optimum for s_i is reached for $p_i(t)$ defined by Equation (9) and $q_{ik}(t)$ defined by Equation (12). If one of these inequalities is not true, then the optimum for s_i is reached on the frontier defined by Equation (7).

In the rest of the article, the game settings (energy needs of the microgrids, energy production and penalties) will be chosen so that we are always in energy shortage, in the sense that Inequality (14) will always be true, and with fair penalties, in the sense of Inequality (13). As a result, the optimal price for s_i is defined by Equation (9) and the optimal orders for s_i are defined by Equation (12).

3.3 Optimization of the energy producers' decision

After substituting $q_{ik}(t)$ and $C_k(t)$ by the expressions found in the previous section in energy producer e_k 's utility as defined in Equation (4), we obtain:

$$\tilde{\pi}_k(t) = \nu_k^e(t) \frac{n-1}{\delta} \left(\sum_{i=1, \dots, n} \frac{L(i)}{\tilde{\gamma}_i} - \sum_{i=1, \dots, n} \left(\frac{L(i)}{\tilde{p}_k(t)} \left(1 - \frac{\tilde{p}_k(t) \tilde{\gamma}_i \delta}{n-1} \right)_+ \right) \right)$$

The only part of this equation depending on $\tilde{p}_k(t)$ has always a negative impact on the profit of the energy producer under the assumption of fair penalties. Indeed, in that case, as raised in the previous section, we have: $L(i) \geq 0$ for all service providers s_i . As a result, to maximize his profit, the energy producer has to choose $\tilde{p}_k(t)$ such that the part depending on $\tilde{p}_k(t)$ in the above equation equals 0. It implies that the term $1 - \frac{\tilde{p}_k(t) \tilde{\gamma}_i \delta}{n-1}$ is inferior to 0 for all $i = 1, \dots, n$. It is equivalent to: $\tilde{p}_k(t) \geq \frac{n-1}{\delta \tilde{\gamma}_i}$. Consequently, the optimal price for the energy producer with fair penalties is defined by:

$$\tilde{p}_k(t) = \frac{n-1}{\delta \min_{i=1, \dots, n} \{\tilde{\gamma}_i\}}$$

In theory, the price could be higher than this value and it would change nothing for the utility of the energy producer. But the energy producer has an incentive to be moderate on his price to avoid competition from other energy producers.

4 Distributed learning game

In this section, we will assume that the consumer needs $\nu_i^s(t)$ and the energy productions $\nu_k^e(t)$ are random individual sequences. This means that the underlying random processes generating the sequences do not necessarily have a probabilistic structure.

They can be quite erratic. Compared with [4] and [8], the producers' available energy cannot be controlled and are not known by the energy producers before the end of the time period. The reason behind this is that the energy comes from renewable sources exclusively.

In the previous section, we have defined the optimal decisions for each agent at time period t . We have proved that these decisions do depend neither on the consumer needs $\nu_i^s(t)$ nor on the energy productions $\nu_k^e(t)$ except for the service providers. To guarantee the optimal operation of the whole system, it is fundamental for the service providers to elaborate efficient learning strategies about consumer needs and energy productions.

To simplify, we will consider a common space \mathcal{X}_e of possible values for the production of each energy producer and a common space \mathcal{X}_s of possible values for the consumption of each service provider's associated microgrid. $\mathcal{X}_e, \mathcal{X}_s \subseteq \mathbb{R}$ are supposed to be of finite dimension i.e., their cardinals $|\mathcal{X}_e|$ and $|\mathcal{X}_s|$ are such that $|\mathcal{X}_e| < +\infty$ and $|\mathcal{X}_s| < +\infty$. In the description of the generic learning algorithm, we will use the notation \mathcal{X} to designate indifferently \mathcal{X}_s or \mathcal{X}_e .

Service providers should optimize their prices and booking quantities at each time period, having no information about the produced energy and the energy needs of the microgrids at this instant. As a result, the game can be considered as having *partial information* [2]. Each service provider s_i has to forecast $\nu_i^s(t)$ and $\nu_k^e(t)$ for all $k = 1, \dots, K$ at each time period in order to optimize his decisions. We will denote by $f_i(X, t)$ the forecast of service provider s_i about the variable X at time period t . We will also use the simplifying notations:

- $f_i(t) = \left\{ f_i(\nu_i^s, t), f_i(\nu_1^e, t), \dots, f_i(\nu_K^e, t) \right\}$ to denote the predictions made by service provider s_i about microgrid \mathcal{M}_i instantaneous needs and about the production of each energy producer e_k , $k = 1, \dots, K$.
- $f(t) = \left\{ f_1(t), \dots, f_n(t) \right\}$ which contains the forecasts of all the service providers.
- $f_{-i}(y, t) = \left\{ f_1(t), \dots, f_{i-1}(t), y, f_{i+1}(t), \dots, f_n(t) \right\}$ which contains the forecasts of all the service providers except s_i which prediction is set equal to y .
- $\nu(t) = \left\{ \nu_1^s(t), \dots, \nu_n^s(t), \nu_1^e(t), \dots, \nu_K^e(t) \right\}$ which contains the microgrid energy needs and the production of each energy producer e_k , $k = 1, \dots, K$.

By substitution of the forecasters in the Stackelberg game solution at equilibrium as obtained in Section 3, we infer the optimal decisions for service provider s_i at each time period t :

$$\begin{aligned}
 p_i(t) &= \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \\
 q_{ik}(t) &= \frac{f_i(\nu_k^e, t)}{\tilde{p}_k} \frac{L(i)}{\tilde{\gamma}_i} \frac{n-1}{\delta}
 \end{aligned}$$

As a result, the utility of service provider s_i at each time period t is:

$$\begin{aligned}
\pi_i(t) &= \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \left(\nu_i^s(t) - \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \right) \\
&- \frac{L(i)}{\tilde{\gamma}_i} \frac{n-1}{\delta} \sum_{k=1, \dots, K} f_i(\nu_k^e, t) - \gamma_i \left(\nu_i^s(t) - \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \right) \\
&- \sum_{k=1, \dots, K} \frac{f_i(\nu_k^e, t) L(i)}{\sum_{j=1, \dots, n} f_j(\nu_k^e, t) L(j)} \nu_k^e(t) \Big)_+ \tag{15}
\end{aligned}$$

4.1 Loss, external regret and Hannan consistency definitions

Provider s_i 's loss is defined as:

$$l_i(f(t), \nu(t)) = \left(\pi_i^0(t) - \pi_i(t) \right)$$

where $\pi_i^0(t)$ corresponds to provider s_i 's utility evaluated in $f_i(\nu_i^s, t) = \nu_i^s(t)$ and $f_i(\nu_k^e, t) = \nu_k^e(t)$, $\forall k = 1, \dots, K$. It means that $\pi_i^0(t)$ contains the utility that provider s_i would have received if his forecasts were perfectly aligned with microgrid \mathcal{M}_i instantaneous needs and with the production of each energy producer.

Having no a priori information about the dynamic evolution of the produced renewable energies and about the energy needs of the microgrids, we assume that everything happens as if the system were in the worst case: Nature and consumers allie together to form a *meta-player* who is supposed to be the most unfavorable to the service providers. It means that the meta-player tries to maximize the sum of the providers' losses. His loss can be expressed as the opposite of the sum of all the providers' losses. Therefore, it takes the form:

$$l(f(t), \nu(t)) = \sum_{i=1, \dots, n} \left(\pi_i(t) - \pi_i^0(t) \right)$$

The agents' external regret over the sequence of time periods $1, \dots, T$, is expressed as the realized difference between the cumulative loss and the loss of the best prediction i.e., pure strategy (in the sense that this prediction minimizes their cumulative loss).

To be more precise, for service provider s_i , it coincides with the difference between s_i 's truly observed cumulative loss and the cumulative loss that would be obtained in case where s_i made the best constant prediction over time interval $[1; T]$. It takes the form:

$$\mathcal{R}_i(T) = \sum_{t=1}^T l_i(f(t), \nu(t)) - \min_{y \in \mathcal{X}_s \times \mathcal{X}_e^K} \sum_{t=1}^T l_i(f_{-i}(y, t), \nu(t))$$

Finally, for the meta-player, the regret coincides with the difference between his cumulative loss and the loss of the constant predictions over $[1; T]$ about the unknown sequences minimizing his cumulative loss or equivalently maximizing the sum of the

providers' losses over the interval. We have:

$$\mathcal{R}(T) = \sum_{t=1}^T l(f^s(t), \nu(t)) - \min_{z \in \mathcal{X}_s^n \times \mathcal{X}_e^K} \sum_{t=1}^T l(f(t), z)$$

The service providers and the meta-player try to determine randomized strategies such that asymptotically their external regrets remain in $o(T)$ where T is the number of time periods which have been played. It means that with probability 1:

$$\lim_{T \rightarrow +\infty} \sup \frac{1}{T} \sum_{t=1}^T \mathcal{R}_i(t) = 0$$

for provider s_i , $i = 1, \dots, n$ and

$$\lim_{T \rightarrow +\infty} \sup \frac{1}{T} \sum_{t=1}^T \mathcal{R}(t) = 0$$

for the meta-player. Forecasters satisfying these inequalities are said *Hannan consistent* [2].

In the following lemma, we prove that it is possible to construct learning strategies for the service providers which minimize their external regret asymptotically.

Lemma 1. *A Hannan consistent learning strategy exists for each service provider s_i .*

Proof of Lemma 1. In our case setting, at the end of each time period, service provider s_i knows the energy quantity bought by his customers \mathcal{M}_i and he can infer $\nu_i^s(t)$ from that quantity. s_i also knows the energy which has been delivered by each energy producer e_k to him. He can infer from that the energy which could have been delivered to him, if he had ordered a different quantity $q_{ik}(t)$, all other providers ordering the same energy quantities. As a result, s_i can calculate his loss for all his possible actions. In [2], it is proved that a Hannan consistent learning strategy always exists when the player can calculate his loss for each possible action at the end of each time period. \square

4.2 Analysis of s_i 's loss function upper bounds

In this subsection, we unfold sequentially results enabling us to obtain upper bounds on the provider s_i 's loss function. We will then study the upper bounds on the meta-player's loss in next subsection. Last subsection will cover collaborative learning strategies.

We start by upper bounding provider s_i 's loss through the identification of upper bounds in Proposition 2. This upper bound is represented as the sum of a loss function depending only on provider s_i 's predictions and on another one relying exclusively on the disagreement between provider s_i and the other providers on the predictions of

producer e_k 's production in renewable energy. Then properties of these upper bounds are studied in Corollary 3.

We introduce:

$$d_{ij}^k(t) = f_i(\nu_k^e, t) - f_j(\nu_k^e, t), \quad i, j = 1, \dots, n, \quad k = 1, \dots, K$$

It is a measure of the *disagreement between provider s_i and provider s_j for $i \neq j$* , in the prediction of e_k 's energy production, at time period t .

Proposition 2. *Provider s_i 's loss can be upper-bounded by a sum of two functions: the first one $l_i^{(1)}(\cdot)$ depending only on his forecasts $f_i(t)$ and the second one $l_i^{(2)}(\cdot)$ depending on his disagreement with the other providers' predictions exclusively. Formally, we have: $\forall i = 1, \dots, n$, $l_i(f(t), \nu(t)) \leq l_i^{(1)}(f_i(t), \nu(t)) + l_i^{(2)}((d_{ij}^k(t))_{j,k}, \nu(t))$.*

Proof of Proposition 2. The proof can be found in Appendix.

We now demonstrate functional properties for provider s_i 's upper bounds which lead us to the following observations concerning the provider's strategic learning behavior.

Corollary 3. *To minimize his loss, provider s_i should:*

- *Forecast the exact real value of the microgrid \mathcal{M}_i energy needs.*
- *Minimize his energy production forecasts and his disagreements with the other providers in the energy production forecasts at the same time.*

Proof of Corollary 3. Judging by the form of function $l_i^{(1)}(\cdot)$ as obtained in the proof of Proposition 2 which is detailed in Appendix, it is linear increasing in $f_i(\nu_k^e, t)$, $\forall k = 1, \dots, K$ since $L(i) \geq 0$ as proved in Section 3.2. Therefore, to reduce his loss, s_i has incentives to choose small values for $f_i(\nu_k^e, t)$, $\forall k = 1, \dots, K$.

Furthermore, the differentiate of $l_i^{(1)}(\cdot)$ in $f_i(\nu_i^s, t)$ equals 0 when $f_i(\nu_i^s, t) = \nu_i^s(t)$. Since $l_i^{(1)}(\cdot)$ is a second order polynomial in $f_i(\nu_i^s, t)$ with a positive first coefficient, the minimum of $l_i^{(1)}(\cdot)$ is reached in $\nu_i^s(t)$. As a result, to reduce his loss, s_i has incentives to choose $f_i(\nu_i^s, t) = \nu_i^s(t)$.

Since the differentiate of $l_i^{(2)}(\cdot)$ in $d_{ij}^k(t)$ is always positive, $l_i^{(2)}(\cdot)$ increases when the disagreement with the other providers $d_{ij}^k(t)$ increases. \square

In the rest of the paper, functions $l_i^{(1)}(y, \nu(t))$ will be denoted the *partial losses* for provider s_i , this latter making predictions $y \in \mathcal{X}_s \times \mathcal{X}_e^K$ while the meta-player chooses $\nu(t) \in \mathcal{X}_s^n \times \mathcal{X}_e^K$. We start by introducing lower and upper bounds on the disagreements between provider s_i and the other providers about the predictions of the energy productions:

$$\underline{D}_{ss}(i) = \min_{t=1, \dots, T} \min_{j \neq i, k} d_{ij}^k(t)$$

and

$$\overline{D}_{ss}(i) = \max_{t=1,\dots,T} \max_{j \neq i,k} d_{ij}^k(t)$$

They contain the extreme disagreement values between the providers, about the estimated energy productions.

Lemma 4. *If provider s_i plays according to a Hannan consistent strategy for his loss upper bound then, there exists an upper bound for the external regret associated with s_i 's partial loss which depends only on the extreme disagreement values between the providers about the estimated energy productions $\underline{D}_{ss}(i)$ and $\overline{D}_{ss}(i)$.*

Proof of Lemma 4. The proof can be found in Appendix.

4.3 Analysis of the sum of service providers loss functions upper bounds

The repetition of the Stackelberg game introduced in Section 3 in a context of partial information can be rewritten by introducing randomization in the strategies. We denote by $d_t(f_i) : \mathcal{X}_s \times \mathcal{X}_e^K \rightarrow [0; 1]$ and $d_t(\nu) : \mathcal{X}_s^n \times \mathcal{X}_e^K \rightarrow [0; 1]$ the randomized strategies for service provider s_i and for the meta-player respectively at time period t . We then have to cope with a repeated learning game. At each time period t , the repeated game timing introduced in Section 3 is updated according to the following rules to incorporate the forecasting tasks of the providers:

- (1) All the service providers s_i , $i = 1, \dots, n$ make their forecasts $f_i(\nu_i^s, t)$, $f_i(\nu_k^e, t)$, $\forall k = 1, \dots, K$ following distributions $d_t(f_i)$ respectively.
- (2) Energy producers reveal their energy prices.
- (3) Service producers reveal their energy orders $q_{ik}(t)$ and their service prices at the same time.
- (4) The meta-player chooses $\nu_i^s(t)$ and $\nu_k^e(t)$, $\forall i = 1, \dots, n$ and $\forall k = 1, \dots, K$ following the distribution $d_t(\nu)$.
- (5) Each service provider s_i obtains his profit $\pi_i(t)$, the demand of \mathcal{M}_i and the energy quantities offered by each service producer e_k .

Service providers update their forecasting strategies $d_t(f_i)$ and the meta-player updates his forecasting strategy $d_t(\nu)$ depending on the value of the expected utilities.

The step corresponding to the generation of unexpected random events resulting in microgrid energy needs and production variations is now controlled by the meta-player whereas the penalty rules introduced in Section 3 remain unchanged.

We define now $\tilde{l}_g(\cdot)$ as the sum of the providers' partial losses:

$$\tilde{l}_g(f(t), \nu(t)) = \sum_{i=1,\dots,n} l_i^{(1)}(f_i(t), \nu(t)) \quad (16)$$

If the repeated learning game reaches a stable state i.e., an equilibrium for the generating density functions, the randomized strategy becomes time invariant for t large enough:

$$\begin{aligned} d_t(f_i) &\rightarrow d(f_i) \\ d_t(\nu) &\rightarrow d(\nu) \end{aligned}$$

We let F_s be the set of all the predictors (i.e., discrete density function set or alternatively, randomized prediction set) for each service provider and F_m the set of all the predictors for the meta-player. It will be used to properly introduce the value of the game. The value of the game where the providers consider their partial losses as utilities is defined as:

$$\tilde{V}_g = \min_{\otimes_{i=1,\dots,n} d(f_i) \in F_s^n} \max_{d(\nu) \in F_m} \tilde{l}_g^E \left(\otimes_{i=1,\dots,n} d(f_i), d(\nu) \right)$$

where $\tilde{l}_g^E(\cdot)$ represents the expectation of function $\tilde{l}_g(\cdot)$ as defined in Equation (16).

Theorem 5. *Assume that all service providers play according to Hannan consistent strategies for their loss upper bound. Then:*

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \tilde{l}_g(f(t), \nu(t)) \leq \tilde{V}_g + \frac{1}{T} \sum_{t=1}^T \sum_{i=1,\dots,n} \psi \left(\underline{D}_{ss}(i), \overline{D}_{ss}(i) \right) \sum_{k=1,\dots,K} \nu_k^e(t)$$

where we have defined the function $\psi(\cdot, \cdot)$ from \mathbb{R}^2 to \mathbb{R} such that:

$$\psi \left(\underline{D}_{ss}(i), \overline{D}_{ss}(i) \right) = \gamma_i L(i) \sum_{k=1,\dots,K} \nu_k^e(t) \left(\frac{1}{g(\underline{D}_{ss}(i))} - \frac{1}{g(\overline{D}_{ss}(i))} \right) \quad (17)$$

with $g(x) = 1 - \frac{x}{\max\{\mathcal{X}_e\}} \mathbf{1}_{x \geq 0} - \frac{x}{\min\{\mathcal{X}_e\}} \mathbf{1}_{x < 0}$.

Proof of Theorem 5. The proof can be found in Appendix.

Corollary 6. *Assume that the meta-player plays according to a Hannan consistent strategy for his loss upper bound. Then:*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \tilde{l}_g(f(t), \nu(t)) \geq \tilde{V}_g - \frac{1}{T} \sum_{t=1}^T \sum_{i=1,\dots,n} \psi \left(\underline{D}_{ss}(i), \overline{D}_{ss}(i) \right) \sum_{k=1,\dots,K} \nu_k^e(t)$$

Proof of Corollary 6. Applying Theorem 5 to the meta-player i.e, by symmetry, considering that the meta-player's loss upper bound is the opposite of the sum over i

of s_i 's loss upper bounds, and using von Neuman-Morgenstern's minimax theorem [9] for \tilde{V}_g , we derive the proposed inequality. \square

We let:

$$l_g(f(t), \nu(t)) = \sum_{i=1, \dots, n} l_i(f(t), \nu(t)) \quad (18)$$

be the sum of the providers' losses. Using the definitions settled in Equations (16) and (18), we derive the following inequality:

$$l_g(f(t), \nu(t)) \leq \tilde{l}_g(f(t), \nu(t)) + \sum_{i=1, \dots, n} l_i^{(2)}\left((d_{ij}^k(t))_{j,k}, \nu(t)\right)$$

where j browses all the values in the set $\{1, \dots, n\}$ and k , all the values in the set $\{1, \dots, K\}$. By substitution in Theorem 5, we obtain the following result:

Corollary 7. *If all service providers play according to a Hannan consistent strategy for their loss upper bounds then, their average loss cannot be larger than:*

$$\begin{aligned} & \tilde{V}_g + \frac{1}{T} \sum_{t=1}^T \sum_{i=1, \dots, n} \psi\left(\underline{D}_{ss}(i), \overline{D}_{ss}(i)\right) \sum_{k=1, \dots, K} \nu_k^e(t) \\ & + \frac{1}{T} \sum_{t=1}^T \sum_{i=1, \dots, n} l_i^{(2)}\left((d_{ij}^k(t))_{j,k}, \nu(t)\right) \end{aligned}$$

whatever strategy is chosen by the meta-player.

The objective of the next subsection is to determine which conditions to impose on the gaps between the providers' predictions, in order to minimize the upper bound of the agents' average loss.

4.4 Collaborative learning strategy

As explained in the previous sections, to optimize their decisions, the service providers need to forecast both their associated microgrid energy needs as well as the production of all the energy producers. Judging by the result highlighted in Corollary 7, to minimize the upper bound of their average loss, providers have no choice but to collaborate to predict the energy productions. However, the providers still predict independently their associated microgrid instantaneous needs since there is no interaction between the microgrids, each of them adapting independently the bought quantity of energy to the time varying price.

Collaboration takes place within coalitions. In cooperative game theory literature, a *coalition* is a group of agents who have incentives to collaborate by sharing resource access, information, etc., in the hope to increase their revenue, knowledge, social welfare

(in case of altruism), etc., compared to the case where they behave non-cooperatively [9], [13]. Adapted to our hierarchical learning context, we define coalitions of agents as follows:

Definition 8. • *A coalition of agents is a group of agents who share their information and align their predictions to a common value.*

- *Agents who belong to the same coalition are said to collaborate.*

At this stage, the objective is to identify conditions on the disagreement levels between the providers about the forecasted energy productions such that the term at the right of \tilde{V}_g defined in Corollary 7, remains as small as possible. Indeed, the smaller is the term defined in Corollary 7, the smaller is the sum of the agents' losses.

Such a strategy would satisfy the following relations, at any time period t :

$$\begin{aligned} \psi\left(\underline{D}_{ss}(i), \overline{D}_{ss}(i)\right) \sum_{k=1, \dots, K} \nu_k^e(t) &= 0, \quad \forall i = 1, \dots, n \\ \Leftrightarrow \underline{D}_{ss}(i) &= \overline{D}_{ss}(i), \quad \forall i = 1, \dots, n \end{aligned}$$

It means that service providers can decrease their total loss by aggregating their predictions about the produced energies $\nu_k^e(t)$, $\forall k = 1, \dots, K$, at any time period t . Providers therefore have incentive to form a grand coalition because it might enable them to decrease their total loss.

By substitution in the second part of the loss upper bounds as introduced in Propositions 2 and detailed analytically in Appendix, we let:

$$\mu_i(t) = l_i^{(2)}\left((d_{ij}^k(t))_{j,k}, \nu(t)\right) \Big|_{d_{ij}^k(t)=0, j=1, \dots, n, k=1, \dots, K}$$

and we obtain: $\mu_i(t) = -\gamma_i L(i) \sum_{k=1, \dots, K} \nu_k^e(t)$ which depends only on the service provider index (i) and on time period t , but not on the providers' forecasts.

Proposition 9. *When the providers collaborate through a grand coalition, the system average loss over time interval $[1; T]$ cannot be larger than $\tilde{V}_g + \frac{1}{T} \sum_{t=1}^T \sum_{i=1, \dots, n} \mu_i(t)$.*

Proof of Proposition 9. By definition:

$$\begin{aligned} l_g(t) &= \sum_{i=1, \dots, n} l_i(t) \\ &\leq \underbrace{\sum_{i=1, \dots, n} l_i^{(1)}(t)}_{\tilde{l}_g(t)} + \sum_{i=1, \dots, n} l_i^{(2)}(t) \\ &\leq \tilde{l}_g(t) + \sum_{i=1, \dots, n} \mu_i(t). \end{aligned}$$

Taking the average of these values over time interval $[1; T]$, the proposition statement is straightforward. \square

5 Simulations

The aim of this section is to explain how the economic model of the hierarchical network described in Section 2 can be applied in practice to take decisions in an uncertain context and then to check that the results derived analytically in Section 4 hold, for a smart grid which structure is defined a priori.

The rest of the section is organized as follows: Subsection 5.1 deals with payoff function estimation for each forecast, Subsection 5.2 elaborates on the update of mixed strategies for each forecast and we discuss in the last part the results that we have obtained on a numerical example.

5.1 Payoff functions

As detailed in Section 4, after the energy prices have been issued by the energy producers, the service providers make their forecasts and choose a quantity to order to each producer and their optimal energy unit price according to the following system:

$$\begin{aligned} p_i(t) &= \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \\ q_{ik}(t) &= \frac{f_i(\nu_k^e, t)}{\bar{p}_k} \frac{L(i)}{\bar{\gamma}_i} \frac{n-1}{\delta} \end{aligned}$$

By substitution in Equation (2), service provider s_i 's utility can then be rewritten as:

$$\begin{aligned} \pi_i(t) &= \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \left(\nu_i^s(t) - \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \right) - \frac{L(i)}{\bar{\gamma}_i} \frac{n-1}{\delta} \\ &\quad \sum_{k=1, \dots, K} f_i(\nu_k^e, t) - \gamma_i \left(\nu_i^s(t) - \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \right) \\ &\quad - \sum_{k=1, \dots, K} \frac{L(i) f_i(\nu_k^e, t)}{\sum_{j=1, \dots, n} L(j) f_j(\nu_k^e, t)} \nu_k^e(t) \end{aligned}$$

Interestingly, it is worth mentioning that the profit of the service provider depends on the forecasts of the other service providers.

Each service provider has $K + 1$ forecasts to do: one for his associated microgrid's consumption and one to evaluate the productions of each of the K energy producers. In principle, it should be necessary to define a randomized strategy on space $\mathcal{X}_s \times \mathcal{X}_e^K$ for each service provider. We recall that a randomized strategy is the classical terminology used in game theory to designate a discrete density function defined over the considered set [9]. This quantity grows very fast with K and as a result each probability in the randomized strategy of forecasts is very small which leads to rounding errors during computation. In order to overcome this issue, we have decided to cut the service providers in smaller entities, each of them making only one forecast at each time

period and to consider that these entities are uncoupled. This trick results in $K + 1$ randomized strategies in the space of forecasts $\mathcal{X}_s \times \mathcal{X}_e^K$ for each service provider.

For a given forecast X , we derive the payoffs for each value $x \in \mathcal{X}$ of the forecast at each time period t by using the utilities of the service providers and keeping only the terms depending on forecast X . This is summarized in the following definition:

Definition 10. *The payoff function associated to forecast X evaluated in $x \in \mathcal{X}$ coincides with the utility of service provider s_i restricted to its terms depending on forecast X solely and evaluated in x .*

For the forecasts of associated microgrid \mathcal{M}_i energy needs, provider s_i 's payoff takes the form:

$$\begin{aligned} H_{f_i(\nu_i^s)}(x, t) &= \frac{x + \gamma_i}{2} \left(\nu_i^s(t) - \frac{x + \gamma_i}{2} \right) - \gamma_i \left(\nu_i^s(t) - \frac{x + \gamma_i}{2} \right. \\ &\quad \left. - \sum_{k=1, \dots, K} \frac{L(i) f_i(\nu_k^e, t)}{\sum_{j=1, \dots, n} L(j) f_j(\nu_k^e, t)} \nu_k^e(t) \right)_+ \end{aligned}$$

Concerning the forecasts of energy producer e_k 's production, provider s_i 's payoff takes the form:

$$\begin{aligned} H_{f_i(\nu_k^e)}(x, t) &= -\frac{L(i)}{\tilde{\gamma}_i} \frac{n-1}{\delta} x - \gamma_i \left(\nu_i^s(t) - \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \right. \\ &\quad \left. - \sum_{l \neq k} \frac{L(i) f_i(\nu_l^e, t)}{\sum_j L(j) f_j(\nu_l^e, t)} \nu_l^e(t) - \frac{L(i)x}{\sum_{j \neq i} L(j) f_j(\nu_k^e, t) + L(i)} \right. \\ &\quad \left. \nu_k^e(t) \right)_+ \end{aligned}$$

As already stated in Section 4, we will also consider that the meta-player is non oblivious and plays so as to minimize the sum of the utilities of the service providers. As for the service providers, we uncouple $\nu_i^s(t)$ and $\nu_k^e(t)$ to improve the computation. More precisely the meta-player's payoffs are:

$$\begin{aligned} H_{\nu_i^s}(x, t) &= \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \left(\frac{f_i(\nu_i^s, t) + \gamma_i}{2} - x \right) + \gamma_i \left(x - \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \right. \\ &\quad \left. - \sum_{k=1, \dots, K} \frac{L(i) f_i(\nu_k^e, t)}{\sum_{j=1, \dots, n} L(j) f_j(\nu_k^e, t)} \nu_k^e(t) \right)_+ \end{aligned}$$

$$\begin{aligned} H_{\nu_k^e}(x, t) &= \sum_{i=1, \dots, n} \gamma_i \left(\nu_i^s(t) - \frac{f_i(\nu_i^s, t) + \gamma_i}{2} - \sum_{l \neq k} \frac{L(i) f_i(\nu_l^e, t)}{\sum_{j=1, \dots, n} L(j) f_j(\nu_l^e, t)} \right. \\ &\quad \left. \nu_l^e(t) - \frac{L(i) f_i(\nu_k^e, t)}{\sum_{j=1, \dots, n} L(j) f_j(\nu_k^e, t)} x \right)_+ \end{aligned}$$

It is very straightforward to adapt the repeated learning game and payoffs considering that the service providers integrate a grand coalition. The grand coalition payoffs take the following forms:

$$\begin{aligned} H_{f_C(\nu_i^s)}(x, t) &= H_{f_i(\nu_i^s)}(x, t) \\ H_{f_C(\nu_k^e)}(x, t) &= - \sum_{i=1, \dots, n} \frac{L(i)}{\tilde{\gamma}_i} \frac{n-1}{\delta} x \end{aligned}$$

Whereas, the meta-player's payoffs become:

$$\begin{aligned} H_{\nu_i^s}(x, t) &= \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \left(\frac{f_i(\nu_i^s, t) + \gamma_i}{2} - x \right) + \gamma_i \left(x - \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \right. \\ &\quad \left. - L(i) \sum_{k=1, \dots, K} \nu_k^e(t) \right)_+ \\ H_{\nu_k^e}(x, t) &= \sum_{i=1, \dots, n} \gamma_i \left(\nu_i^s(t) - \frac{f_i(\nu_i^s, t) + \gamma_i}{2} - L(i) \left(\sum_{l \neq k} \nu_l^e(t) + x \right) \right)_+ \end{aligned}$$

5.2 Updates of forecasting strategies

In the rest of the paper, we consider two types of updates for the forecasting randomized strategies $d_t(X)$ at each time period based on the exponential forecaster for signed games: one based on the external regret and the other based on the internal regret [2]. We assume that this is a signed game because the range of values of payoff function $H_X(\cdot)$ might include a neighborhood of 0.

We let:

$$\begin{aligned} \mathcal{V}_t &= \sum_{s=1}^t \text{Var} \left(H_X(X_s, s) \right) \\ &= \sum_{s=1}^t \mathbb{E} \left[\left(H_X(X_s, s) - \mathbb{E}[H_X(X_s, s)] \right)^2 \right] \end{aligned}$$

be the sum of the variances associated with the random variable $H_X(X_t, t)$ which is the payoff for forecaster X at time period t assuming that the forecast at time period t has been set to X_t , under the mixed strategy X which is defined over space \mathcal{X} . Using the exponential forecaster for signed games with external regret means that the randomized strategy is updated according to the algorithm described below.

External Regret Learning Algorithm: Updating of the Exponential Forecaster

Initialization. For $t = 0$, we set: $w_0(x) = \frac{1}{|\mathcal{X}|}$, $\forall x \in \mathcal{X}$.

Step 1 to T . The updating rules are the following:

$$\begin{aligned}
d_{t+1}(x) &= \frac{w_{t+1}(x)}{\sum_{x \in \mathcal{X}} w_{t+1}(x)}, \forall x \in \mathcal{X} \\
w_{t+1}(x) &= \exp\left(\eta_{t+1} \sum_{s=1}^t H_X(x, s)\right) \\
&= d_t(x)^{\frac{\eta_{t+1}}{\eta_t}} \exp\left(\eta_{t+1} H_X(x, t)\right), \forall x \in \mathcal{X} \\
\eta_{t+1} &= \min\left\{\frac{1}{2 \max\{|H_X(\cdot)|\}}; \sqrt{\frac{2(\sqrt{2}-1)}{e-2}} \sqrt{\frac{\ln|\mathcal{X}|}{\mathcal{V}_t}}\right\} \\
\mathcal{V}_t &= \mathcal{V}_{t-1} + \text{Var}\left(H_X(X_t, t)\right)
\end{aligned}$$

For the internal regret, it is similar but with $d_t(\cdot) = \sum_{i \neq j} d_t^{i \rightarrow j}(\cdot) \Delta_{(i,j)}(t)$ where $d_t^{i \rightarrow j}(\cdot)$ is the modified forecasting strategy obtained when the forecaster predicts j each time he would have predicted i and $\Delta_{(i,j)}(t) = \frac{\omega_{(i,j)}(t)}{\sum_{k \neq l} \omega_{(k,l)}(t)}$ with:
 $\omega_{(i,j)}(t) = \exp\left(\eta_t \sum_{s=1}^{t-1} \sum_{x \in \mathcal{X}} d_s(x) H_X(x, s)\right)$.

We see that we need to compute the maximum of the absolute value of the payoff function $|H_X(\cdot)|$ for all forecasts X to run a simulation of the game. This maximum is reached for $x = \min\{\mathcal{X}\}$ or $x = \max\{\mathcal{X}\}$ for all payoff functions except for $H_{f_i(\nu_i^s)}(\cdot)$ because their differentiate with respect to x is never equal to 0. For $H_{f_i(\nu_i^s)}(\cdot)$, the differentiate equals 0 if, and only if, $f_i(\nu_i^s, t) = \nu_i^s(t)$, so the maximum of $|H_X(\cdot)|$ is reached either for $x = \min\{\mathcal{X}\}$ or $x = \max\{\mathcal{X}\}$ or $x = \nu_i^s(t)$.

5.3 Results

For our numerical illustration, we have chosen $n = 3$ and $K = 2$. We have also used $\gamma_1 = \gamma_2 = \gamma_3 = 0.9$ and $\tilde{\gamma}_1 = 0.5$, $\tilde{\gamma}_2 = 0.4$, $\tilde{\gamma}_3 = 0.6$ and $\mathcal{X}_e = [1; 2]$, $\mathcal{X}_s = [5; 8]$ which ensure that the $L(i)$, $i = 1, 2, 3$ remain positive and that Inequality (14) is always true.

In the following pictures, we compare the cumulative regret of each player to the cumulative regret of the same player who would have forecasted the best value at each time period in terms of payoffs. More precisely, we display:

$$\frac{1}{t} \sum_{s=1}^t \sum_{X \in F} \left(H_X(X_s, s) - \max_x (H_X(x, s)) \right)$$

where F is the generic set of forecasts made by the service provider or the meta-player or the coalition considered.

We start by comparing the cumulative internal and external regrets in the case of full competition between service providers in Figures 2 and 3.

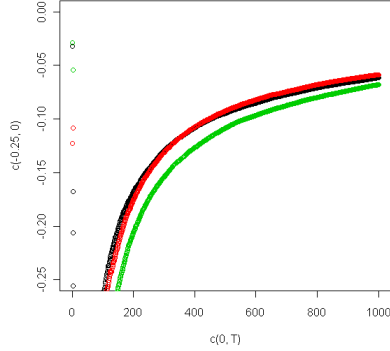


Figure 2: Difference between the best achievable cumulative regret and the one obtained with the internal regret minimization algorithm under full competition.

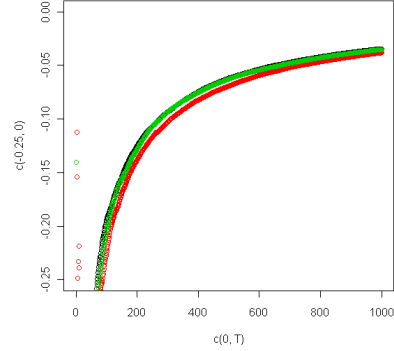


Figure 3: Difference between the best achievable cumulative regret and the one obtained with the external regret minimization algorithm under full competition.

The service providers are in black for s_1 , green for s_2 and red for s_3 . We can see that in all cases, the difference between regrets converge toward 0 which means that the cumulative payoff obtained at the end of the game following the exponential forecaster strategy is close to the best possible cumulative payoff. This is in coherence with the theoretical result for the internal regret but is better than what we could expect for the external regret which means that we are in a game setting which performs well for regret based learning. We also remark that the algorithm converges faster for the external regret compared to the internal regret.

We compare these graphs with the graphs obtained when service providers integrate a grand coalition in Figures 4 and 5.

Again, we observe that the differences between the best achievable regrets and those obtained converge toward 0. The rate of convergence under cooperative learning seems higher than in the non-cooperative case. In addition, we observe that after 400 time periods the sum of differences between regrets under collaborative learning is close to -0.2 whereas the sum of differences between regrets is close to -0.26 in the full competition case. This is in coherence with the theory which says that collaborative learning is better.

6 Conclusion

In this article, we focus on the identification of emerging economic organizations while agents interacting through a hierarchical network, are learning to predict as accurately as possible hidden individual sequences modeling variations in renewable energy productions and in microgrid energy needs. Agents belong to one of the 3 categories: producers, providers and end users. We want to prove whether collaboration between

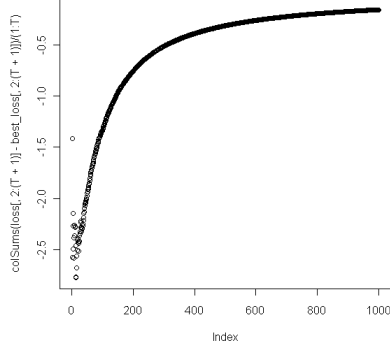


Figure 4: Difference between the best achievable cumulative regret and the one obtained with the internal regret minimization algorithm for the coalition of service providers.

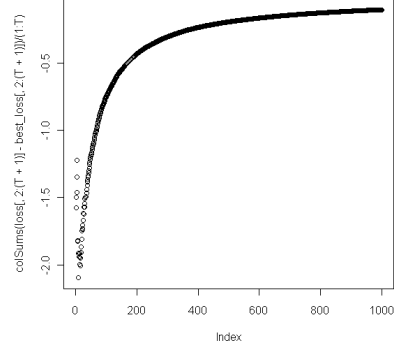


Figure 5: Difference between the best achievable cumulative regret and the one obtained with the external regret minimization algorithm for the coalition of service providers.

the agents through the learning process, can increase the system performances. Collaboration can occur when coalitions are forming. Agents are said to belong to a common coalition when they share their information and align their forecasts on a common prediction.

Over each time period, we have optimized analytically the agents' decisions in prices and traded quantities of energy by solving a Stackelberg game in a complete information context. However, in practice, these decisions should depend on variables which are unobserved at the beginning of each time period such as the energy productions and the microgrid energy needs. The behavior of these processes might be quite erratic which explains why they might be so complex to predict. Additionally, we make the assumption that providers are penalized in case where they cannot satisfy the attached microgrid entire demand and that producers are penalized in case where they cannot deliver the quantity of energy booked at the beginning of the time period by the providers. Consequently, it is necessary for producers and providers to develop efficient distributed learning algorithms.

We prove that there exists prices for the producers that provide them some guarantees to avoid penalties. Therefore, only providers make forecasts. We demonstrate that under external regret minimization, collaborative learning when service providers integrate a grand coalition, enables them to minimize their average loss. This theoretical result is illustrated on a 2 producers, 3 providers smart grid. We observe that the rates of convergence under collaborative learning based on regret minimization is higher than in the non-cooperative case.

In extensions, it might be interesting to determine conditions guaranteeing the sys-

tem resilience. Resilient systems are organizations sharing the capacity to automatically reconfigure themselves in case of shocks or attacks. This is of particular interest for smart grids which are highly sensitive to attacks arising from malicious external agents that might cause unbearable interruptions in the consumers' services. The capacity for the system to reconfigure itself might be directly correlated with its capacity to anticipate external threats by stabilizing itself in economic organizations enabling the sharing of information and collaborative learning.

Appendix

Proof of Proposition 2

By definition of provider s_i 's loss and thanks to Equation (15), we have:

$$\begin{aligned} l_i(f(t), \nu(t)) &= \pi_i^0(t) - \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \left(\nu_i^s(t) - \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \right) \\ &+ \frac{L(i)}{\tilde{\gamma}_i} \frac{n-1}{\delta} \sum_{k=1, \dots, K} f_i(\nu_k^e, t) + \gamma_i \left(\nu_i^s(t) - \frac{f_i(\nu_i^s, t) + \gamma_i}{2} \right) \\ &- \sum_{k=1, \dots, K} \frac{f_i(\nu_k^e, t) L(i)}{\sum_{j=1, \dots, n} f_j(\nu_k^e, t) L(j)} \nu_k^e(t) \end{aligned}$$

For $l = 1, \dots, K + 1$, the l -th component of vector $y \in \mathcal{X}_s \times \mathcal{X}_e^K$ will be denoted: $y(l)$. If we let:

$$\begin{aligned} l_i^{(1)}(y, \nu(t)) &= \pi_i^0(t) - \frac{y(1) + \gamma_i}{2} \left(\nu_i^s(t) - \frac{y(1) + \gamma_i}{2} \right) \\ &+ \frac{L(i)}{\tilde{\gamma}_i} \frac{n-1}{\delta} \sum_{k=1, \dots, K} y(k+1) + \gamma_i \left(\nu_i^s(t) - \frac{y(1) + \gamma_i}{2} \right) \end{aligned}$$

and

$$\begin{aligned} l_i^{(2)}\left((d_{ij}^k(t))_{j,k}, \nu(t)\right) &= -\gamma_i L(i) \sum_{k=1, \dots, K} \nu_k^e(t) \left[\sum_{j=1, \dots, n} L(j) \left(1 \right. \right. \\ &\left. \left. - \frac{d_{ij}^k(t)}{\min\{\mathcal{X}_e\}} \mathbf{1}_{d_{ij}^k(t) \leq 0} - \frac{d_{ij}^k(t)}{\max\{\mathcal{X}_e\}} \mathbf{1}_{d_{ij}^k(t) \geq 0} \right) \right]^{-1} \end{aligned}$$

Using the fact that we are in energy shortage in the sense of Inequality (14) and that $\min\{\mathcal{X}_e\} \leq f_i(\nu_k^e, t) \leq \max\{\mathcal{X}_e\}$, $\forall k = 1, \dots, K$, we can check easily that provider s_i 's loss function can be upper-bounded by $l_i^{(1)}(f_i(t), \nu(t)) + l_i^{(2)}\left((d_{ij}^k(t))_{j,k}, \nu(t)\right)$. It is the sum of two functions: the first one depending only on s_i 's forecasts, and the second one depending only on his disagreement with the other providers' forecasts. \square

Proof of Lemma 4

Suppose that provider s_i plays according to a Hannan consistent strategy according to his loss upper bound i.e., $l_i^{(1)}(f_i(t), \nu(t)) + l_i^{(2)}((d_{ij}^k)_{j,k}(t), \nu(t))$. This means that:

$$\limsup_{T \rightarrow +\infty} \frac{1}{T} \left[\sum_{t=1}^T l_i^{(1)}(f_i(t), \nu(t)) + \sum_{t=1}^T l_i^{(2)}((d_{ij}^k(t))_{j,k}, \nu(t)) \right] - \min_{y \in \mathcal{X}_s \times \mathcal{X}_e^K} \left(\sum_{t=1}^T l_i^{(1)}(y, \nu(t)) + \sum_{t=1}^T l_i^{(2)}((d_{ij}^k(y, t))_{j,k}, \nu(t)) \right) \leq 0 \quad (19)$$

where $d_{ij}^k(y, t)$ contains the disagreement between provider s_i and all the other providers when s_i makes the prediction y at time period t without any change in the predictions of other providers.

In Section 4, we have introduced lower and upper bounds on the disagreements between provider s_i and the other providers about the forecasts of the energy productions $\underline{D}_{ss}(i)$ and $\overline{D}_{ss}(i)$. According to Corollary 3, $l_i^{(2)}(\cdot)$ being increasing in $d_{ij}^k(t)$, it is possible to provide lower and upper bounds for the function by evaluating it in $\underline{D}_{ss}(i)$ and $\overline{D}_{ss}(i)$ respectively. The lower bound is:

$$b_l(i, t) = -\frac{\gamma_i L(i)}{g(\underline{D}_{ss}(i))} \sum_{k=1, \dots, K} \nu_k^e(t)$$

Whereas, the upper bound takes the form:

$$b_u(i, t) = -\frac{\gamma_i L(i)}{g(\overline{D}_{ss}(i))} \sum_{k=1, \dots, K} \nu_k^e(t)$$

If Inequality (19) is checked, then the following inequality holds:

$$\limsup_{T \rightarrow +\infty} \frac{1}{T} \left[\sum_{t=1}^T l_i^{(1)}(f_i(t), \nu(t)) + \sum_{t=1}^T b_l(i, t) - \min_{y \in \mathcal{X}_s \times \mathcal{X}_e^K} \left(\sum_{t=1}^T l_i^{(1)}(y, \nu(t)) - \sum_{t=1}^T b_u(i, t) \right) \right] \leq 0$$

This last inequality provides an upper bound for the external regret associated with provider s_i 's partial loss. □

Proof of Theorem 5

With the proposed expression of ψ , the upper bound of the external regret evaluated in provider s_i 's partial loss becomes:

$$\begin{aligned} & \limsup_{T \rightarrow +\infty} \frac{1}{T} \left[\sum_{t=1}^T l_i^{(1)}(f_i(t), \nu(t)) - \min_{y_i \in \mathcal{X}} \sum_{t=1}^T l_i^{(1)}(y_i, \nu(t)) \right] \\ & \leq \frac{1}{T} \psi(\underline{D}_{ss}(i), \overline{D}_{ss}(i)) \sum_{t=1}^T \sum_{k=1, \dots, K} \nu_k^e(t) \end{aligned} \quad (20)$$

Summing Inequality (20) over all $i = 1, \dots, n$, the external regret evaluated in the sum of the providers' partial losses, becomes:

$$\begin{aligned} & \limsup_{T \rightarrow +\infty} \frac{1}{T} \left[\sum_{t=1}^T \tilde{l}_g(f(t), \nu(t)) - \min_{f(\cdot)} \sum_{t=1}^T \tilde{l}_g(f(\cdot), \nu(t)) \right] \\ & \leq \frac{1}{T} \sum_{i=1, \dots, n} \psi(\underline{D}_{ss}(i), \overline{D}_{ss}(i)) \sum_{t=1}^T \sum_{k=1, \dots, K} \nu_k^e(t) \end{aligned}$$

In addition:

$$\min_{f(\cdot)} \frac{1}{T} \sum_{t=1}^T \tilde{l}_g(f(\cdot), \nu(t)) = \min_{\otimes_{i=1, \dots, n} d(f_i) \in F_s^n} \frac{1}{T} \sum_{t=1}^T \tilde{l}_g^E(\otimes_{i=1, \dots, n} d(f_i), \nu(t))$$

where \tilde{l}_g^E represents the expectation of function \tilde{l}_g .

We assume that each service provider makes his forecasts independently of the other service providers. Then $\tilde{l}_g^E(\cdot, \nu(t))$ is linear in $\otimes_{i=1, \dots, n} d(f_i)$. As a result, its minimum over the simplex of probability vectors is reached in one of the corners of the simplex.

Let: $d_T(z) = \frac{1}{T} \sum_{t=1}^T \mathbf{1}_{\{\nu(t)=z\}}$ be the marginal empirical frequency of play evaluated in prediction $z \in \mathcal{X}_s^n \times \mathcal{X}_e^K$.

Then:

$$\begin{aligned} & \min_{\otimes_{i=1, \dots, n} d(f_i) \in F_s^n} \frac{1}{T} \sum_{t=1}^T \tilde{l}_g^E(\otimes_i d(f_i), \nu(t)) \\ & = \min_{\otimes_{i=1, \dots, n} d(f_i) \in F_s^n} \sum_{z \in \mathcal{X}_s^n \times \mathcal{X}_e^K} d_T(z) \tilde{l}_g^E(\otimes_i d(f_i), z) \\ & = \min_{\otimes_{i=1, \dots, n} d(f_i) \in F_s^n} \tilde{l}_g^E(\otimes_{i=1, \dots, n} d(f_i), d_T(\cdot)) \\ & \leq \underbrace{\max_{d(\nu) \in F_m} \min_{\otimes_{i=1, \dots, n} d(f_i) \in F_s^n} \tilde{l}_g^E(\otimes_{i=1, \dots, n} d(f_i), d(\nu))}_{\tilde{V}_g} \end{aligned}$$

□

References

- [1] Bubeck S., Online Optimization, Lecture Notes, Princeton university, Department of Operations Research and Financial Engineering, 2012
- [2] Cesa-Bianchi N., Lugosi G., Prediction, Learning, And Games, Cambridge University Press, 2006
- [3] Fudenberg D., Levine D. K., The Theory of Learning in Games, the MIT Press, 1998
- [4] Le Cadre H., Bedo J.-S., Distributed Learning in Hierarchical Networks, Alternative Energies and Atomic Energy Commission Research Paper, available at <https://sites.google.com/site/helenelecadre/research>
- [5] Le Cadre H., Potarusov R., Auliac C., Energy Demand Prediction: A Partial Information Game Approach, in proc. EEVC 2012
- [6] Li N., Marden J. R., Decoupling coupled constraints through utility design, Discussion paper, Department of ECEE, university of Colorado, Boulder, 2011
- [7] Li H., Zhang W., QoS routing in smart grids, in proc. IEEE global communication conference, 2010
- [8] Marden J. R., Young H. P., Pao L. Y., Achieving Pareto Optimality Through Distributed Learning, Oxford Economics discussion paper n°557, 2011
- [9] Myerson R., Game Theory: An Analysis of Conflict, Harvard university press, 2006
- [10] Nguyen T., Vojnović M., Weighted Proportional Allocation, in proc. of ACM Sigmetrics 2011
- [11] Pradelski B. R., Young H. P., Learning efficient Nash equilibria in distributed systems, discussion paper, Department of Economics, university of Oxford, 2010
- [12] Saad W., Han Z., Debbah M., Hjørungnes A., Başar T., Coalitional Game Theory for Communication Networks: A Tutorial, IEEE Signal Processing Magazine, Special Issue on Game Theory, vol.26, pp.77–97, 2009
- [13] Saad W., Han Z., Poor V. H., Coalitional Game Theory for Cooperative Micro-Grid Distribution Networks, 2-nd IEEE International Workshop on Smart Grid Communications, 2011
- [14] Shapley L. S., Stochastic games, in proc. of the National Academy of Sciences of the United States of America, vol.39, pp.1095–1100, 1953
- [15] Talluri K. T., van Ryzin G. J., The Theory and Practice of Revenue Management, International Series in Operations Research & Management Science, vol.68, 2005
- [16] Young H. P., Learning by Trial and Error, Games and Economic Behavior, vol.65, pp.626–643, 2009