Information Management in the Smart Grid: A Learning Game Approach

Hélène Le Cadre, Jean-Sébastien Bedo

Information Management in the Smart Grid: A Learning Game Approach

Hélène Le Cadre
MINES ParisTech,
Centre for Applied Mathematics,
Sophia Antipolis, France
helene.le_cadre@mines-paristech.fr

Jean-Sébastien Bedo
Orange/France Télécom,
Place d'Alleray, 75015 Paris, France
jeansebastien.bedo@orange.com

Abstract

In this article, the smart grid is modeled as a decentralized and hierarchical network, made of three categories of agents: producers, providers and microgrids. To optimize their decisions concerning the energy prices and the traded quantities of energy, the agents need to forecast the energy productions and the demand of the microgrids. The biases resulting from the decentralized learning might create imbalances between demand and supply, leading to penalties for the providers and for the producers. We determine analytically prices that provide to the producers a guarantee to avoid such penalties, reporting all the risk on the providers. Additionally, we prove that collaborative learning, through a grand coalition of providers where information is shared and forecasts aligned on a single value, minimizes their average risk. Simulations, run on a toy network, lead us to observe that the convergence rates of the collaborative learning strategy are clearly superior to rates resulting from distributed learning, using external and internal regret minimization.

Keywords: Distributed Learning; Information; Regret; Learning Game Theory

1 Introduction

In Europe, and in France especially, traditional electrical networks rely on nuclear based energies [8], which are non renewable energy sources. With such energies, the production level can be adapted by the plant operator who alternates openings and closings and optimizes the duration of the switches between both modes. The objective then, is to adapt the production level so as to meet the uncertain demand level. We have built a first model in [11], where two learning strategies based on tit for tat and fictitious play are used to adapt the production level to meet the demand level. For renewable energies, the production level can only be partially controlled, for instance, by lowering the wind turbine speed [15]. Renewable energy integration in the electrical network requires the deployment of smart Information and Communication Technologies (ICTs), to supervise the grid operation [5]. Indeed, renewable energy production is highly unpredictable since it depends on uncontrolable exogenous factors like level of wind, sunshine, swell, etc. [17]. Furthermore, the new active role of the end users, who can become energy producers and adapt dynamically their consumption while falling into a multitude of microgrids [8], [12], [20], [21], dramatically increase the volume of the exchanged data flows. ICTs appear as a means to retrieve the most salient information from this big amount of data and to train forecasters to provide efficient predictions about the renewable energy production and about the microgrid energy demand. These predictions will then be used as inputs to optimize the smart grid operation.

In practice, it is increasingly apparent that current forecasting technology cannot properly handle extreme situations corresponding to either extreme weather phenomena or critical periods for power system operation. For example, forecasting methods used to predict wind power were mostly designed to provide single value forecasts of expected productions. Only recently, probabilistic methods have been introduced to provide estimations of the entire distribution of future productions [1]. In such methods, forecasts may take the form of either quantile estimations or density estimations [6], [9]. Learning based on regret minimization, as described in [4], belongs to this latter category. This class of methods is particulary efficient [10], [13]. It provides to the forecaster a density function which associates a weight to each possible outputs. The density function is updated by merging informations from various sources' reports. As a result, these methods are more robust to extreme events and appear as particularly well suited to model erratic processes such as renewable energy production.

In the framework of the smart grid, learning is performed in a decentralized manner since each agent primarly learns the hidden information using his own observations. The existing literature on distributed learning primarly focuses on distributed learning algorithms that are suitable for implementation in large scale engineering systems [14], [19]. The results mainly concentrate on a specific class of games, called games of potential [22]. 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 exist 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 [15]. 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 energy orders, 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 [19]. 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 [15].

Most collaborative mechanisms studied in the literature lead to price or quality of service alignment. To our knowledge the impact of collaboration, through information sharing and forecast alignment while prices are individually determined, on the underlying system performance, has not been studied so far. Of course, collaboration might not emerge due to the agents' natural incentives to cheat and to deviate from the cooperative equilibrium and also, most frequently, due to the regulator's intervention. There is a number of well-understood reasons why regulator often does not allow horizontal collaboration: if providers are allowed to collaborate, they might cooperate to raise the price i.e., reducing quantity below the efficient baseline, and create market power [7]. Alternatively, providers might cooperate to reduce quality of service. Courts punish explicit agreements whose objectives are clearly to decrease the competition [2]. In this article, we will answer the following questions:

- How will the biases, introduced by the errors made by the agents in their predictions, affect the agents' average risk?
- Will collaborative learning improve the smart grid wide performance and should therefore be encouraged by the regulator?

The article is organized as follows. In Section 2, we introduce the economic basis of our model, the agents, their utility and their optimization program. Then the complete information Stackelberg game is solved in Section 3, proceeding by backward induction. We derive analytically the optimal prices and energy orders 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 operation in Section 5 and an illustration of the previously derived theoretical results is provided for a toy network.

2 The model

The number of agents interacting in the smart grid is large. In this article, we model the smart grid as a three 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 energy production and demand variations. We detail the three categories of agents and the repeated game which captures the interactions between them in Subsection 2.1. Then, we define each agent optimization program in Subsection 2.2.

2.1 Description of the agents

We model the smart grid through three categories of agents: the microgrids, the providers and the energy producers. The end users generate some energy demand, and fulfill it either by buying energy to a provider or by finding alternatives (solar panels, more efficient appliances, etc.). Each provider buys energy to several energy producers and resell it to the end users. Each energy producer produces and sells energy to all providers.

We assume that each end user contracts with only one provider and does not churn from one 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 end users attached to a single provider can be seen as an individual microgrid, as defined in [12], [20], [21] and recalled in the Introduction. We denote by $(s_i)_i$, with i varying between 1 and n, the i-th provider and by $\mathcal{M}_i$ the corresponding group of end users. The energy producers are denoted $(e_k)_k$ with k varying between 1 and K. The energy producers can be associated with nuclear plants, photovoltaic park managers, wind farm administrators, etc. In this article, we assume that the energy producer cannot influence directly the energy he produces at each time period. This assumption holds well if we look at renewable energy sources like 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 produced energy without any lever for the energy producer [1].

We model the interplay between all the agents through a repeated game. At each time period t, the following game is played:

Basic Game Description $G(t)$

---

(1) The energy producers e_k communicate their unitary prices $\tilde{p}_k(t) > 0$ to the providers. The energy prices are fixed independently and simultaneously by each energy producer so as to maximize his profit.

(2) The providers s_i place energy quantity orders to energy producers: the quantity ordered by s_i to e_k is denoted by $q_{ik}(t)$. The providers s_i communicate their prices $p_i(t) > 0$ for one energy unit to their microgrid. The orders and the energy prices are fixed independently and simultaneously by each provider so as to maximize his profit.

(3) Microgrid $\mathcal{M}_i$ demand reaches $\nu_i^s(t)$ energy units for the time period. It decides to find alternatives for $a_i(t)$ energy units and buys the rest $\Big(\nu_i^s(t) - a_i(t)\Big)_+$ to provider s_i. The quantity of alternatives is chosen so as to minimize the total cost of energy for $\mathcal{M}_i$.

At each time period, the production of energy producer e_k reaches $\nu_k^e(t)$ energy units. He then delivers $\alpha_{ki}(t)\nu_k^e(t)$ energy units to provider s_i where $\alpha_{ki}(t) \geq 0$ denotes the proportion of his production that producer e_k allocates to provider s_i, with the normalization constraint: $\sum_{i=1,\ldots,n} \alpha_{ki}(t) = 1$. This proportion is defined depending on quantity orders received by e_k from all providers. The sum of all quantity orders may exceed $\nu_k^e(t)$ and so the quantity of energy received by each provider may be inferior to his quantity order.

Penalties are undertaken by both providers and energy producers if they cannot satisfy the entire demand of their consumers. Provider s_i incurs a cost $\gamma_i > 0$ per missing energy unit for his microgrid, measured a posteriori. It is paid to the electricity transmission network operator. In France, the electricity transmission network operator has defined some rules to give incentives to the agents to become balance operators [23]. According to these rules, a positive energy balance is paid the spot price and a negative balance is paid the price defined through the adjustement mechanism. This latter is implemented by the electricity transmission operator who compensates the negative balances to ensure the electrical network reliability. In our article, the spot

[1]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 [11], [15].

price is set to zero and the adjustement price is different for each provider. The price discrimination is justified by the fact that depending on its geographic location, a negative energy balance can be easily corrected in densely interconnected areas whereas it is much more difficult in isolated ones due to the high cost of electricity transmission. As a result, γ_i is higher for providers serving isolated locations than over densely interconnected areas. Producer e_k incurs a cost $\tilde{\gamma}_i > 0$ per missing energy unit for provider s_i, measured a posteriori. It is paid to the regulator of the capacity market that should be implemented to balance supply and demand in the smart grid [24]. Indeed, to guarantee the reliability of the capacity market, it might be necessary to implement a feedback mechanism where the regulator compensates the negative energy balances of the producers by investing himself in capacity [24]. The costs of these investments would be recovered from the penalties imposed to the producers.

2.2 Optimization program for each agent

In this subsection, we describe the decision variables and the utilities for each category of agents. The optimization program for each agent is presented using a mathematical formulation.

2.2.1 Programs of the microgrids

The only decision variable for microgrid $\mathcal{M}_i$ is the quantity of energy that it decides to get from alternative sources: $a_i(t)$. We assume that the microgrid has no lever to influence its random demand: $\nu_i^s(t)$.

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

$$p_i(t)\Big(\nu_i^s(t) - a_i(t)\Big) + c\Big(a_i(t)\Big) \tag{1}$$

We assume that the cost is quadratic in the alternatives i.e., $c\Big(a_i(t)\Big) = \frac{a_i(t)^2}{2}$. This assumption is not restrictive and constants or more complicated cost functions can be introduced. The main advantage of this choice is that it is generic enough and allows to solve a large part of the game analytically.

Mircrogrid $\mathcal{M}_i$ chooses $a_i(t)$ in order to minimize its total energy cost. Therefore, its optimization program is of the form: $\min_{a_i(t)\geq 0}\Big\{p_i(t)\Big(\nu_i^s(t)-a_i(t)\Big)+c\Big(a_i(t)\Big)\Big\}$. Its decision depends on the energy price $p_i(t)$ fixed by provider s_i.

2.2.2 Programs of the providers

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

Throughout the article, we will use the notation: $x_+ = \max\{x; 0\}$ to denote the positive part of the real number x.

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

$$\begin{aligned}\pi_i(t) &= p_i(t)\Big(\nu_i^s(t) - a_i(t)\Big) - \sum_{k=1,\ldots,K} q_{ik}(t)\tilde{p}_k(t) - \gamma_i\Big(\nu_i^s(t) - a_i(t) \\ &- \sum_{k=1,\ldots,K} \alpha_{ki}(t)\nu_k^e(t)\Big)_+ \end{aligned} \tag{2}$$

Provider s_i chooses his energy unit price and his energy orders toward energy producers so that $\pi_i(t)$, as defined in Equation (2), is maximized. His optimization program takes the form: $\max_{p_i(t)>0,(q_{ik}(t))_k\in\mathbb{R}_+^K}\Big\{\pi_i(t)\Big\}$.

2.2.3 Programs of the 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 providers.

The utility of energy producer e_k at time period t equals:

$$\tilde{\pi}_k(t) = \tilde{p}_k(t)\sum_{i=1,\ldots,n} q_{ik}(t) - \sum_{i=1,\ldots,n} \tilde{\gamma}_i\Big(q_{ik}(t) - \alpha_{ki}(t)\nu_k^e(t)\Big)_+$$

To define the sharing coefficients $\Big(\alpha_{ki}(t)\Big)_i$, 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 [21] to weighted energy orders with penalty coefficients as weights. Such a resource sharing mechanism has already been introduced by Nguyen and Vojnović, in [18]. This means that between two providers with identical energy orders, 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 serving isolated areas where failure of electricity supply may be critical. We set:

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

where $C_k(t) = \sum_{j=1,\ldots,n} \tilde{\gamma}_j q_{jk}(t)$. Using Equation (3), energy producer e_k's utility at time period t can be rewritten:

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

Energy producer e_k chooses his energy unit price so that $\tilde{\pi}_k(t)$, as defined in Equation (4), is maximized. His optimization program is of the form: $\max_{\tilde{p}_k(t)>0}\Big\{\tilde{\pi}_k(t)\Big\}$.

3 Complete information game resolution

The game setting described in Subsection 2.1 implies that in the relation producers-providers, producers appear as leaders whereas providers are followers. Identically, in the relation providers-microgrids, providers appear as leaders whereas microgrids are mere followers. Under such a setting, the game is called a Stackelberg game and, as usual, it should be solved using backward induction [12], [16].

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

3.1 Optimization of the microgrids' decision

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

$$a_i(t) = p_i(t) \tag{5}$$

3.2 Optimization of the providers' decisions

To find his optimal price and energy orders, 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.

Case 1: the energy production fulfills the energy demand of the microgrid

It is the case when:

$$\nu_i^s(t) - p_i(t) < \sum_{k=1,\ldots,K} \alpha_{ki}(t)\nu_k^e(t) \tag{6}$$

Then differentiating the 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, $\alpha_{ki}(t)$ will tend toward zero. This implies in turn that s_i will tend to break the inequality defining Case 1 in Inequality (6) and we will always fall on the frontier between Case 1 and Case 2. The frontier between these two cases is defined by the equation:

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

Case 2: the energy production does not fulfill the energy demand of the microgrid

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

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

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 the system of Equations (8), we conclude that the differentiates equal 0 when:

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

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

On the 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:

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

If s_i anticipates that the other providers will make the same optimization program, replicating Equation (11) for the n providers and summing them all, results in the following equality: $C_k(t) = nC_k(t) - \frac{\tilde{p}_k(t)C_k(t)^2}{\nu_k^e(t)} \sum_{j=1,\dots,n} \frac{1}{\gamma_j \tilde{\gamma}_j}$ 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: $C_k(t) = \frac{\nu_k^e(t)}{\tilde{p}_k(t)} \frac{n-1}{\delta}$ where $\delta = \displaystyle\sum_{j=1,\dots,n} \frac{1}{\gamma_j \tilde{\gamma}_j}$. By replacing $C_k(t)$ in Equation (11), we obtain the energy orders for which the differentiates of $\pi_i(t)$ equals 0:

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

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:

$$\gamma_i \tilde{\gamma}_i \geq \frac{n-1}{\delta} \tag{13}$$

This inequality means that the penalties related to s_i are close to the penalties related to the other providers. Indeed, if all penalties are equal to γ, then $\delta = \frac{n}{\gamma^2}$ and Inequality (13) is true for all 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) for s_1 becomes $n \geq (n-1)^2$ which is false as soon as $n > 2$. In this case, s_1 would not buy any energy to the producers and so would be out of the game.

Second, by replacing the energy orders, defined by Equation (12), in Equation (3), we obtain $\alpha_{ki}(t) = \frac{L(i)}{\displaystyle\sum_{j=1,\dots,n} L(j)} = L(i)$ meaning that the total energy delivered to microgrid $\mathcal{M}_i$, attached to s_i, is: $\sum_{k=1,\ldots,K} \alpha_{ki}(t)\nu_k^e(t) = L(i) \sum_{k=1,\ldots,K} \nu_k^e(t)$. As a result, the price and energy orders for which the differentiates of $\pi_i(t)$ equal 0 verify the inequality defining Case 2 if, and only if:

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

This inequality states that the total production of energy by energy producers should not be too large compared to the demand of the microgrid. If it would not be the case, the over provisioning situation may probably get the most expensive producer out of the game.

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 parameters and random events (demands of the microgrids, energy productions 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:

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

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 providers $(s_i)_i$. 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 $i = 1, ..., n$. 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 should satisfy:

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

4 Distributed learning game

In this section, we assume that the microgrids demands $\left(\nu_i^s(t)\right)_i$ and the energy productions $\left(\nu_k^e(t)\right)_k$ are random individual sequences. As explained in the Introduction, this means that the underlying random processes generating the sequences do not necessarily have a probabilistic structure. They can be quite erratic [3], [4].

In the previous section, we defined the optimal decisions for each agent at time period t. We proved that these decisions do depend neither on the microgrids demands $\nu_i^s(t)$ nor on the energy productions $\nu_k^e(t)$ except for the providers. To guarantee the optimal system wide operation, it is fundamental for the providers to elaborate efficient learning strategies about the demand of the microgrids and about the productions of the energy producers. The risk associated with this learning task will be measured by the provider's loss. It will be defined in Subsection 4.1.

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 demand of a 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 $|\mathcal{X}_s| < +\infty$.

Providers should optimize their prices and ordered quantities at each time period, having no information about the produced energy and the demands of the microgrids at this instant. As a result, the game can be considered as having partial information [4]. Each provider s_i has to forecast $\nu_i^s(t)$ and $\nu_k^e(t)$ for all $k = 1, ..., K$, at each time period, in order to optimize his decisions. We will denote by $f_i(X, t)$ the forecast of provider s_i about the variable X at time period t. We will also use the simplifying notations:

- $f_i(t) = \Big\{f_i(\nu_i^s, t), f_i(\nu_1^e, t), ..., f_i(\nu_K^e, t)\Big\}$ to denote the predictions made by provider s_i about microgrid $\mathcal{M}_i$ demand and about the production of each energy producer $e_k,\ k = 1, ..., K$.

- $f(t) = \Big\{f_1(t), ..., f_n(t)\Big\}$ which contains the forecasts of all the providers.

- $f_{-i}(y, t) = \Big\{f_1(t), ..., f_{i-1}(t), y, f_{i+1}(t), ..., f_n(t)\Big\}$ which contains the forecasts of all the providers except s_i which prediction is set equal to y.

- $\nu(t) = \Big\{\nu_1^s(t), ..., \nu_n^s(t), \nu_1^e(t), ..., \nu_K^e(t)\Big\}$ which contains the microgrids' demand and the production of each energy producer $e_k,\ k = 1, ..., K$.

By substitution of the forecasters in the Stackelberg game solution at equilibrium as obtained in Section 3, we infer the optimal decisions for provider s_i at each time period t: $p_i(t) = \frac{f_i(\nu_i^s,t)+\gamma_i}{2}$ and $q_{ik}(t) = \frac{f_i(\nu_k^e,t)}{\tilde{p}_k(t)}\frac{L(i)}{\tilde{\gamma}_i}\frac{n-1}{\delta}$. As a result, the utility of provider s_i at each time period t is:

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

In this section, the game will be repeated over a finite time horizon $0 < T < +\infty$.

4.1 Learning risk measure definition and first observations

As already mentioned, the provider's risk, associated with the learning task, will be measured by his resulting loss.

For any provider $s_i, i = 1, ..., n$, his loss is defined as: $l_i\Big(f(t), \nu(t)\Big) = \Big(\pi_i^0(t) - \pi_i(t)\Big)$ 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 $f_i(\nu_k^e, t) = \nu_k^e(t)$ for any $k = 1, ..., K$. 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 $\mathcal{M}_i$ demand and with the production of each energy producer.

We start by upper bounding provider s_i's loss as the sum of a loss function depending only on provider s_i's predictions and on another one relying exclusively on the disagrement between provider s_i and the other providers on the predictions of producer e_k's renewable energy production. We introduce:

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

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

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

Proof of Proposition 1. The proof can be found in Appendix.

In the rest of the paper, functions $l_i^{(1)}\Big(y, \nu(t)\Big)$ will be denoted the partial losses for provider s_i, this latter making predictions $y \in \mathcal{X}_s \times \mathcal{X}_e^K$.

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 2. *To minimize his loss, provider s_i should:*

- *Have no bias in his forecast of microgrid $\mathcal{M}_i$ demand*
- *Minimize his energy production forecasts while reducing his disagreements with the other providers*

Proof of Corollary 2. Judging by the form of function $l_i^{(1)}(.)$, as obtained in the proof of Proposition 1 which is detailed in Appendix, it is linear increasing in $f_i(\nu_k^e, t),\ \forall k = 1, ..., K$ 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 $f_i(\nu_k^e, t),\ \forall k = 1, ..., K$.

Furthermore, the differentiate of $l_i^{(1)}(.)$ in $f_i(\nu_i^s, t)$ equals 0 when $f_i(\nu_i^s, t) = \nu_i^s(t)$. Since $l_i^{(1)}(.)$ is a second order polynomial in $f_i(\nu_i^s, t)$ with a positive first coefficient, the minimum of $l_i^{(1)}(.)$ 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 $l_i^{(2)}(.)$ in $d_{ij}^k(t)$ is always positive, $l_i^{(2)}(.)$ increases when the disagreement with the other providers $d_{ij}^k(t)$ increases. □

4.2 Optimal learning strategies for each provider

The external regret over the sequence of time periods $1, ..., T$, is the difference between the observed cumulative loss and the cumulative loss of the best constant prediction i.e., pure strategy. To be more precise, for provider s_i, it takes the form: $\mathcal{R}_i(T) = \sum_{t=1}^{T} l_i\Big(f(t), \nu(t)\Big) - \min_{y \in \mathcal{X}_s \times \mathcal{X}_e^K} \sum_{t=1}^{T} l_i\Big(f_{-i}(y,t), \nu(t)\Big)$. We will consider that the learning strategy of s_i is optimal if asymptotically his external regret remains in $o(T)$ where T is the number of time periods which have been played. It means that with probability 1: $\limsup_{T \to +\infty} \frac{1}{T} \sum_{t=1}^{T} \mathcal{R}_i(t) = 0$. Forecasters satisfying these inequalities are said Hannan consistent [4].

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

Lemma 3. *A Hannan consistent learning strategy exists for each provider s_i.*

Proof of Lemma 3. In our case setting, at the end of each time period, provider s_i knows the energy quantity bought by the microgrid $\mathcal{M}_i$ 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 [4], it is proved that a Hannan consistent learning strategy always exists when the player can compute his loss for each possible action at the end of each time period. □

Having no a priori information about the dynamic evolution of the produced renewable energies and about the microgrids' demand, we assume that everything happens as if the system were in the worst case: Nature and microgrids allie together to form a meta-player who is supposed to be the most unfavorable to the 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: $l\Big(f(t), \nu(t)\Big) = \sum_{i=1,\dots,n} \Big(\pi_i(t) - \pi_i^0(t)\Big)$. Similarly to providers, the meta-player will try to keep its external regret $\mathcal{R}(t)$ in $o(T)$.

We introduce 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 providers loss functions upper bounds

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

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

We let F_s be the set of all the predictors (i.e., discrete density function set or alternatively, randomized prediction set) for each 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: $\tilde{V}_g = \min_{\otimes_{i=1,\dots,n} d(f_i)\in F_s^n} \max_{d(\nu)\in F_m} \tilde{l}_g^E\Big(\otimes_{i=1,\dots,n} d(f_i), d(\nu)\Big)$ where $\tilde{l}_g^E(.)$ represents the expectation of function $\tilde{l}_g(.)$ as defined in Equation (16).

To simplify the analytical derivation of the following theorem, which is detailed in the Appendix, we define the function $\psi(.,.)$ from $\mathbb{R}^2$ to $\mathbb{R}$ such that:

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

where $g(x) = 1 - \frac{x_+}{\max\{\mathcal{X}_e\}} + \frac{(-x)_+}{\min\{\mathcal{X}_e\}}$.

Theorem 5. *Assume that all providers play according to Hannan consistent strategies for their loss upper bound then:* $\limsup_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \tilde{l}_g\Big(f(t),\nu(t)\Big) \le \tilde{V}_g + \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1,\dots,n}\psi\Big(\underline{D}_{ss}(i),\overline{D}_{ss}(i)\Big)\sum_{k=1,\dots,K}\nu_k^e(t)$.

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:* $\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \tilde{l}_g\Big(f(t),\nu(t)\Big) \ge \tilde{V}_g - \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1,\dots,n}\psi\Big(\underline{D}_{ss}(i),\overline{D}_{ss}(i)\Big)\sum_{k=1,\dots,K}\nu_k^e(t)$.

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 [16] for $\tilde{V}_g$, we derive the proposed inequality.

□

We let:

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

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

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

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 providers play according to a Hannan consistent strategy for their loss upper bounds then, their average loss cannot be larger than:*

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

whatever strategy is chosen by the meta-player.

4.4 Collaborative learning strategy

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 [2], [16], [21]. Adapted to our learning context, we define coalitions of agents as follows:

Definition 8.
- *A coalition of providers is a group of providers who collaborate to learn the hidden energy productions* $\left(\nu_k^e(t)\right)_k$.
- *The grand coalition contains all the providers involved in the learning task i.e.,* $\{s_i\}_{i=1,\dots,n}$.
- *Cooperation takes place within the coalition when its members share their information and align their predictions on a common value.*

Shared information concerns only energy productions. Indeed, each provider predict independently its microgrid demand and it has no impact on other providers.

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 upper bound of the sum of the agents' losses.

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

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

It means that providers can decrease the upper bound of their average loss by coordinating their predictions about the produced energies $\left(\nu_k^e(t)\right)_k$, 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 Proposition 1 and detailed analytically in Appendix, we obtain:

$$l_i^{(2)}\Big((d_{ij}^k(t))_{j,k},\nu(t)\Big)|_{d_{ij}^k(t)=0,\ j=1,\dots,n,k=1,\dots,K} = -\gamma_i L(i) \sum_{k=1,\dots,K} \nu_k^e(t) \qquad (19)$$

It depends only on the provider index (i) and on time period t, but not on the providers' forecasts.

Proposition 9. *If the providers cooperate through a grand coalition and play Hannan consistent strategies, the system average loss over time interval* $[1;T]$ *cannot be larger than:* $\tilde{V}_g - \sum_{i=1,\dots,n} \gamma_i L(i) \frac{1}{T}\sum_{t=1}^{T} \sum_{k=1,\dots,K} \nu_k^e(t)$.

Proof of Proposition 9. By definition: $l_g(t) = \sum_{i=1,\dots,n} l_i(t) \leq \underbrace{\sum_{i=1,\dots,n} l_i^{(1)}(t)}_{\tilde{l}_g(t)} + \sum_{i=1,\dots,n} l_i^{(2)}(t) = \tilde{l}_g(t) - \gamma_i L(i) \sum_{k=1,\dots,K} \nu_k^e(t)$. 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

At each time period, each provider has $K+1$ forecasts to do: one for his microgrid demand and one to evaluate the productions of each of the K energy producers. As a result, each provider should define a randomized strategy on the space $\mathcal{X}_s \times \mathcal{X}_e^K$. We recall that a randomized strategy is the classical terminology used in game theory to name a discrete density function defined over the considered set [16]. The size of the set 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 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 provider.

For a given forecast X, we derive the payoffs for each value $x \in \mathcal{X}$ ($\mathcal{X} = \mathcal{X}_s$ for energy demands and $\mathcal{X} = \mathcal{X}_e$ for energy productions) of the forecast at each time period t by using the utilities of the 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 provider s_i restricted to its terms depending on forecast X solely and evaluated in x.*

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

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

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

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

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 providers. As for the providers, we uncouple $\nu_i^s(t)$ and $\nu_k^e(t)$ to improve the computation. More precisely the meta-player's payoffs are:

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

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

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

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

Whereas, the meta-player's payoffs become:

$$
\begin{aligned}
H_{\nu_i^s}(x,t) &= \frac{f_i(\nu_i^s,t)+\gamma_i}{2}\Big(\frac{f_i(\nu_i^s,t)+\gamma_i}{2} - x\Big) + \gamma_i\Big(x - \frac{f_i(\nu_i^s,t)+\gamma_i}{2} \\
&- L(i)\sum_{k=1,\ldots,K}\nu_k^e(t)\Big)_+
\end{aligned}
$$

$$
H_{\nu_k^e}(x,t) = \sum_{i=1,\ldots,n}\gamma_i\Big(\nu_i^s(t) - \frac{f_i(\nu_i^s,t)+\gamma_i}{2} - L(i)(\sum_{l\neq k}\nu_l^e(t)+x)\Big)_+
$$

5.2 Updates of forecasting strategies

In the rest of the paper, we consider two types of updates for the forecasting randomized strategies based on the exponential forecaster for signed games: one based on the external regret and the other based on the internal regret [4]. We assume that this is a signed game because the range of values of payoff function $H_X(.)$ might include a neighborhood of 0.

We let: $\mathcal{V}_t = \sum_{s=1}^t Var\Big(H_X(X_s,s)\Big)$
$= \sum_{s=1}^t \mathbb{E}\Big[\Big(H_X(X_s,s)) - \mathbb{E}[H_X(X_s,s)]\Big)^2\Big]$ be the sum of the variances associated with the random variable $H_X(X_t,t)$ under the mixed strategy $d_t(X)$ which is defined over space $\mathcal{X}$. Using the exponential forecaster for signed games with external regret means that the mixed 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:

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

For the internal regret, it is similar but with $d_t(.) = \sum_{i \neq j} d_t^{i \to j}(.)\Delta_{(i,j)}(t)$ where $d_t^{i \to j}(.)$ 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 $|H_X(.)|$ 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 $H_{f_i(\nu_i^s)}(.)$ because their differentiate with respect to x is never equal to 0. For $H_{f_i(\nu_i^s)}(.)$, the differentiate equals 0 if, and only if, $f_i(\nu_i^s, t) = \nu_i^s(t)$, so the maximum of $|H_X(.)|$ 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:

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

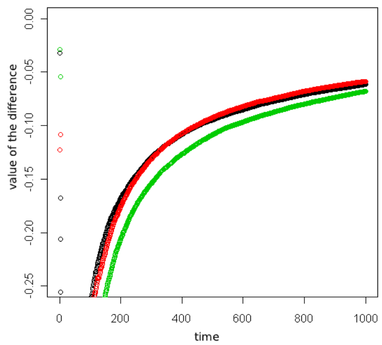
where F is the generic set of forecasts made by the provider or the meta-player or the coalition considered.

We start by comparing the cumulative internal and external regrets in the case of full competition between providers in Figures 2 (a) and 2 (b).

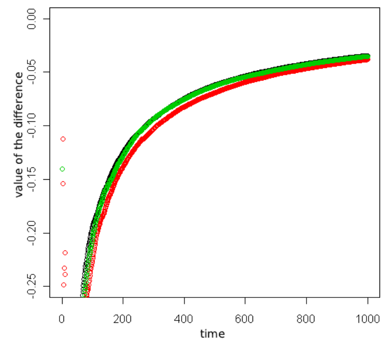
The providers are in black for s_1, green for s_2 and red for s_3. We can see that in all cases, the differences between regrets converge toward 0 which means that the cumulative payoff obtained at the end of the game following the exponential forecaster strategy is close to the best possible cumulative 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 providers integrate a grand coalition in Figures 3 (a) and 3 (b). .

Again, we observe that the differences between the best achievable regrets and those obtained using the learning algorithm converge toward 0. The rate of convergence under cooperative learning seems higher than in the non-cooperative case where the providers perform decentralized learning. In addition, we observe that after 400 time

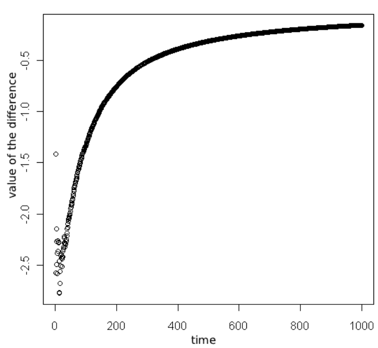


(a)

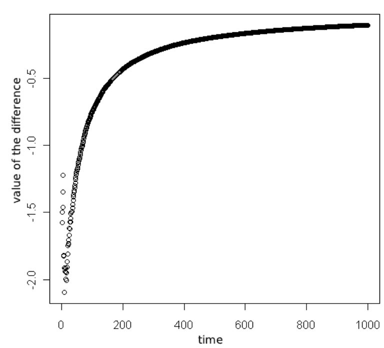


(b)

Figure 1: Difference between the best achievable cumulative regret and the one obtained with the internal regret minimization algorithm in (a) and with the external regret minimization algorithm in (b) under full competition.



(a)



(b)

Figure 2: Difference between the best achievable cumulative regret and the one obtained with the internal regret minimization algorithm in (a) and with the external regret minimization algorithm in (b) for the grand coalition of providers.

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 studied a model of renewable energy production in which producers, providers and microgrids are organized in a hierarchical network. Renewable energy productions were modeled by random individual sequences which need not to have a probabilistic structure. This extraordinarily general demand and supply structure allows to take into account exogenous events. As a result, it is more robust to extreme events and appears as particularly well suited to model quite erratic processes such as renewable energy production. We determined analytically the energy prices enabling the producers to avoid the penalties that the capacity market regulator threatens to apply in case where the providers' orders would not entirely be satisfied. All the risk was then reported on the providers. Additionally, we proved that these latter can minimize their average risk by sharing information and aligning their forecasts. These theoretical results were illustrated on a toy network: we observe that the rates of convergence under collaborative learning through a grand coalition was higher than under decentralized learning, using regret minimization as performance criterion.

An area of improvement concerns the design of the penalties paid to the regulator who compensates the negative energy balances. Is it possible to design more generic mechanisms? Could rules be adapted to guarantee the market opening and avoid speculations, like capacity retention or under investment in the means of production?

Appendix

Proof of Proposition 1

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

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

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

$y(l)$. If we let:

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

and

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

Using the fact that we are in energy shortage in the sense of Inequality (14) and that $\min\{\mathcal{X}_e\} \le f_i(\nu_k^e, t) \le \max\{\mathcal{X}_e\},\ \forall k = 1, ..., K$, we can check easily that provider s_i's loss function can be upper-bounded by $l_i^{(1)}\Big(f_i(t), \nu(t)\Big) + l_i^{(2)}\Big((d_{ij}^k(t))_{j,k}, \nu(t)\Big)$. 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. □

Proof of Lemma 4

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

$$\begin{aligned} \lim\sup_{T\to+\infty} \quad & \frac{1}{T}\Big[\sum_{t=1}^T l_i^{(1)}\Big(f_i(t), \nu(t)\Big) + \sum_{t=1}^T l_i^{(2)}\Big((d_{ij}^k(t))_{j,k}, \nu(t)\Big) \\ -\min_{y\in\mathcal{X}_s\times\mathcal{X}_e^K} \quad & \Big(\sum_{t=1}^T l_i^{(1)}\Big(y, \nu(t)\Big) + \sum_{t=1}^T l_i^{(2)}\Big((d_{ij}^k(y,t))_{j,k}, \nu(t)\Big)\Big)\Big] \le 0 \quad (20) \end{aligned}$$

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 the 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 2, $l_i^{(2)}(.)$ 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:

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

Whereas, the upper bound takes the form:

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

If Inequality (20) is checked, then the following inequality holds:

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

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:

$$\begin{aligned} &\lim \sup_{T\to+\infty} \frac{1}{T}\Big[\sum_{t=1}^{T} l_i^{(1)}\Big(f_i(t),\nu(t)\Big) - \min_{y_i\in\mathcal{X}} \sum_{t=1}^{T} l_i^{(1)}\Big(y_i,\nu(t)\Big)\Big] \\ \leq\quad & \frac{1}{T}\psi\Big(\underline{D}_{ss}(i),\overline{D}_{ss}(i)\Big)\sum_{t=1}^{T}\sum_{k=1,\dots,K}\nu_k^e(t) \end{aligned} \tag{21}$$

Summing Inequality (21) over all $i = 1, ..., n$, the external regret evaluated in the sum of the providers' partial losses, becomes:

$$\begin{aligned} &\lim \sup_{T\to+\infty} \frac{1}{T}\Big[\sum_{t=1}^{T} \tilde{l}_g\Big(f(t),\nu(t)\Big) - \min_{f(.)} \sum_{t=1}^{T} \tilde{l}_g\Big(f(.),\nu(t)\Big)\Big] \\ \leq\quad & \frac{1}{T}\sum_{i=1,\dots,n}\psi\Big(\underline{D}_{ss}(i),\overline{D}_{ss}(i)\Big)\sum_{t=1}^{T}\sum_{k=1,\dots,K}\nu_k^e(t) \end{aligned}$$

In addition: $\min_{f(.)} \frac{1}{T}\sum_{t=1}^{T} \tilde{l}_g\Big(f(.),\nu(t)\Big) = \min_{\otimes_i d(f_i)\in F_s^n} \frac{1}{T}\sum_{t=1}^{T} \tilde{l}_g^E\Big(\otimes_i d(f_i),\nu(t)\Big)$ where $\tilde{l}_g^E$ represents the expectation of function $\tilde{l}_g$. We assume that each provider makes his forecasts independently of the other providers. Then $\tilde{l}_g^E\Big(.,\nu(t)\Big)$ 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 $z \in \mathcal{X}_s^n \times \mathcal{X}_e^K$. Finally,

we obtain: $\min_{\otimes_i d(f_i) \in F_s^n} \frac{1}{T} \sum_{t=1}^{T} \tilde{l}_g^E \Big(\otimes_i d(f_i), \nu(t) \Big)$
$= \min_{\otimes_i d(f_i) \in F_s^n} \sum_{z \in \mathcal{X}_s^n \times \mathcal{X}_e^K} d_T(z) \tilde{l}_g^E \Big(\otimes_i d(f_i), z \Big)$
$= \min_{\otimes_i d(f_i) \in F_s^n} \tilde{l}_g^E \Big(\otimes_i d(f_i), d_T(.) \Big) \leq \underbrace{\max_{d(\nu) \in F_m} \min_{\otimes_i d(f_i) \in F_s^n} \tilde{l}_g^E \Big(\otimes_i d(f_i), d(\nu) \Big)}_{\tilde{V}_g}.$

□

References

[1] Bossavy A., Girard R., Kariniotakis G., Forecasting ramps of wind power production with numerical weather prediciton ensembles, Wind Energy, vol.16, Issue 1, published online feb. 28, 2012

[2] Bourreau M., Coalition, Industrial organization course, IREN Master, 2010

[3] Bubeck S., Online Optimization, Lecture Notes, University of Princeton, Department of Operations Research and Financial Engineering, 2012

[4] Cesa-Bianchi N., Lugosi G., Prediction, Learning, And Games, Cambridge University Press, 2006

[5] Clastres C., Smart grids: Another step towards competition, energy security and climate change objectives, Energy Policy, vol.39, pp.$5399 - -5408$, 2011

[6] Costa A., Crespo A., Navarro J., Lizcano G., Madsen H., Feitosa E., A review on the young history of the wind power short-term prediction, Renewable and Sustainable Energy Reviews, vol.12, Issue 6, pp.$1725 - -1744$, 2008

[7] Creti A., Fabra A., Aupply security and short-run capacity markets for electricity, Energy Economics, vol.29, pp.$259 - -276$, 2007

[8] de Ladoucette P., Chevalier J.-M., The electricity of the future: a mondial challenge, Economica, 2010

[9] Giebel G., Kariniotakis G., Brownsword R., The state of the art on short-term wind power prediction: a literature overview, Working Paper, ANEMOS EU project, 2003

[10] Le Cadre H., Auliac C., Energy Demand Prediction in a Charge Station: A Comparison of Statistical Learning Approaches, in proc. of the 2-nd European Electric Vehicle Congress, EEVC 2012

[11] Le Cadre H., Bedo J.-S., Distributed Learning in Hierarchical Networks, in proc. of the 6-th International Conference on Performance evaluation Methodologies and Tools, ValueTools 2012

[12] Le Cadre H., Mercier D., Is Energy Storage an Economic Opportunity for the Eco-Neighborhood?, NETNOMICS: Economic Research and Electronic Networking, Springer, DOI:$10.1007/s11066-013-9075-7$, 2013

[13] Le Cadre H., Potarusov R., Auliac C., Energy Demand Prediction: A Partial Information Game Approach, in proc. of the 1-st European Electric Vehicle Congress, EEVC 2011

[14] Li N., Marden J. R., Decoupling coupled constraints through utility design, Working Paper, University of Colorado, Department of ECEE, 2011

[15] Marden J. R., Young H. P., Pao L. Y., Achieving Pareto Optimality Through Distributed Learning, Working Paper, University of Oxford, Department of Economics, 2011

[16] Myerson R., Game Theory: An Analysis of Conflict, Harvard University press, 2006

[17] Nair J., Adlakha S., Wierman A., Energy Procurement Strategies in the Presence of Intermittent Sources, Working Paper, California Institute of Technology, 2013

[18] Nguyen T., Vojnović M., Weighted Proportional Allocation, in proc. of ACM Sigmetrics 2011

[19] Pradelski B. R., Young H. P., Learning efficient Nash equilibria in distributed systems, Working Paper, University of Oxford, Department of Economics, 2010

[20] Saad W., Han Z., Debbah M., Hjorrungnes 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

[21] Saad W., Han Z., Poor V. H., Coalitional Game Theory for Cooperative Micro-Grid Distribution Networks, in proc. of the 2-nd IEEE International Workshop on Smart Grid Communications, 2011

[22] 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

[23] Becoming a upstream-downstream balance operator: RTE rules, http://clients.rte-france.com/, 2012

[24] Nome law, http://www.cre.ft/dossiers/la-loi-nome, 2012