



HAL
open science

Towards Forecast Markets For Enhanced Peer-to-Peer Electricity Trading

Ilia Shilov, H el ene Le Cadre, Ana BuŐi c, Anibal Sanjab, Pierre Pinson

► **To cite this version:**

Ilia Shilov, H el ene Le Cadre, Ana BuŐi c, Anibal Sanjab, Pierre Pinson. Towards Forecast Markets For Enhanced Peer-to-Peer Electricity Trading. IEEE SmartGridComm, IEEE, Oct 2023, Glasgow, United Kingdom. 10.1109/smartgridcomm57358.2023.10333930 . hal-04344943

HAL Id: hal-04344943

<https://inria.hal.science/hal-04344943>

Submitted on 14 Dec 2023

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 destin ee au d ep ot et  a la diffusion de documents scientifiques de niveau recherche, publi es ou non,  emanant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv es.



Distributed under a Creative Commons Attribution 4.0 International License

Towards Forecast Markets For Enhanced Peer-to-Peer Electricity Trading

Iliia Shilov^{1,3,*}

Hélène Le Cadre²

Ana Bušić¹

Anibal Sanjab³

Pierre Pinson⁴

Abstract—This paper examines the impact of the coupling between an introduced data market, in which agents can purchase a forecast of their renewable energy sources' generation levels to improve their estimation quality, and a peer-to-peer (P2P) electricity market, enabling prosumers to trade energy in a decentralized manner with their peers, amidst the growing trend of decentralization and uncertainty of renewable energy sources (RES) in electricity markets. The study formulates the P2P trading as a generalized Nash equilibrium problem and identifies conditions for achieving a maximized efficiency of the peer-to-peer electricity market, one of which is prosumers' participation in the forecast market. Along these lines, the analysis demonstrates that prosumers have incentives to participate in the forecast market and outlines the conditions, considering the case in which the forecasts are given in the form of a Gaussian distribution. Numerical examples using Pecan Street data demonstrate the theoretical findings and provide illustrations for the general case, as well as highlight the mutual benefits of market coupling for forecast sellers and electricity market agents.

Index Terms—Mechanism Design, Decentralized Electricity Market, Peer-to-peer Market, Forecast Market

I. INTRODUCTION

The rise of decentralization in electricity markets is driven by the liberalization of the energy sector, an increase in renewable energy sources (RES), and the growing prominence of prosumers - individuals who can both consume and produce energy [1], [2]. In the peer-to-peer (P2P) electricity markets, prosumers negotiate energy transactions, aiming to minimize costs while navigating a complex system of individual and trading reciprocity constraints, and uncertainties of renewable energy production.

In this respect, a large body of literature discusses the management of uncertainty in these markets, employing strategies such as multistage stochastic bidding, flexibility activation, and forward contracts. However, the challenge in decentralized electricity markets is the management of uncertainties from variable renewable energy resources, which necessitates accurate forecasting and knowledge of probability distributions for decision-making [3], [4].

In this study, we present a two-stage model of the P2P electricity market (day-ahead and real-time markets), where agents may purchase (sell) electricity in both stages to guarantee an adequate supply to meet their demand. In the literature, these markets are commonly modeled as a variation of stochastic

inventory models [3], [2]. Such models typically require full knowledge of the probability distribution, which can be difficult to attain due to limited historical data or subjective forecasting techniques. The decision maker must select a probability distribution to use as an input, either relying on their forecasts or those provided by another source, which will affect the computed order quantity [3].

Therefore, forecasting is crucial for making informed decisions and planning, which has rendered this problem widely studied by researchers and practitioners. In [5], the authors present a comprehensive review of the field and its advancements. In the literature dedicated to stochastic inventory models, the effect of the random distribution update was investigated in [3] while in [6] authors investigated the effects of partial information about the random variable distribution (e.g., mean, variance, symmetry, unimodality) available to the decision maker and how updated information affects agents' costs.

Possibility to update the information naturally leads to the concept of information (or data) markets [7]. Prediction (forecast) markets have generated a wide research, see e.g. [8] for a detailed overview. Such markets can be used to aggregate and disperse information into efficient forecasts of uncertain future events. Such a forecast improvement by combining or utilizing more data from various sources was shown in multiple studies, as, e.g., in [9] for energy applications and [10] for forecast competition study analysis. Improvement of wind power prediction accuracy utilizing the spatiotemporal relationship was considered in [11]. In our application, we focus on the forecast market for the improvement of buyers' forecasts in a probabilistic format.

In this study, we make the first step in the connection between the forecast market and the peer-to-peer electricity market. To the best of our knowledge, this is the first work in this direction. We model the electricity market as a game-theoretic model in a form of generalized Nash equilibrium problem (GNPE) [12], [13]. We then adopt a model for the forecast market which fairly rewards forecasters based on the accuracy and their contribution to the client's utility improvement [4], [14].

The key contributions of our study include demonstrating that the impact of forecast updates on a prosumer's outcome can be evaluated independently, allowing us to internalize the utility brought about by the forecast update. Furthermore, we illustrate that the efficiency of the electricity P2P market can be achieved if prosumers participate in the forecast market. We show the coupling to be individually rational for the Gaussian distribution-based forecasts, alongside with intuition for the general case. Lastly, using a Pecan Street dataset, we show that connecting the forecast and electricity markets benefits both

¹ Inria Paris, DI ENS, CNRS, PSL University, France

² Université de Lille, Inria, CNRS, Centrale Lille, UMR 9189 - CRISTAL, France

³ VITO/EnergyVille, Thorpark 8310, Genk, Belgium

⁴ Imperial College London, Dyson School of Design Engineering, London, UK

* Corresponding author, email: ilia.shilov@inria.fr

prosumers and forecast sellers if their forecast enables improved prediction of RES-based generation.

II. COUPLED MARKETS MODEL

A. Agents and Markets

This section introduces two models for peer-to-peer electricity and forecast markets and their coupling framework. We study a two-stage (day-ahead and real-time) model for the electricity market, where prosumers trade in the day-ahead market based on renewable energy generation forecasts. The model involves electricity trading between prosumers and backup retailers in a network represented by the set \mathcal{N} of N agents, located on nodes of an undirected graph $\mathcal{G} := (\mathcal{N}, E)$, where $E \subseteq \mathcal{N} \times \mathcal{N}$ is the set of links between the agents. We aim to design decentralized markets based on network-defined trading relationships, focusing on financial aspects and excluding physical distribution. The real-time market allows agents to observe actual generation and adjust unsettled quantities, representing a stochastic inventory model variant, complexified by inclusion of the peer-to-peer trading. The quality of agents' decisions in such models, and subsequently, their costs, depend on the quality of their forecasts.

In our framework, agents can purchase a forecast before electricity market trading. To do so, they submit their base forecast to the forecast market operator (MO) and receive aggregated forecasts from sellers $i \in I$. We note that the set I does not necessarily coincide with the set \mathcal{N} , but we do not put any restriction on that. Post electricity market clearing, the MO distributes part of the agent's forecast purchase utility among I based on individual contributions. The electricity and forecast market models are introduced next.

B. Peer-to-Peer Electricity Market

We employ a two-settlement electricity market design consisting of day-ahead and balancing (real-time) markets. We assume the presence of a backup retailer from whom the community can purchase energy both in day-ahead (hereafter, referred to as first stage) and in real-time (hereafter, referred to as second stage). Therefore, we fix the buying (b) and selling (s) prices for first (or day-ahead p_{da}) and second (or real time p_{rt}) stages, such that $p^{rt,b} > p^{da,b} > p^{da,s} > p^{rt,s}$. The community is seen as a price-taker in the electricity market, hence making the prices exogenous to this problem.

Let $\Gamma_i \subseteq \mathcal{N}$ denote the set of neighbors of agent i , which reflects the agents with whom she wants to trade. We denote the trade between agent i and $j \in \Gamma_i$ as q_{ij} (limited with upper-limit κ_{ij}), where q_{ij} is the amount of power i purchases from (or sells to) j if $q_{ij} \geq 0$ ($q_{ij} \leq 0$) and impose a bilateral trading reciprocity constraint $q_{ij} + q_{ji} = 0$. Trading cost term is presented in the cost function as $\sum_{j \in \Gamma_i} c_{ij} q_{ij}$, where parameters $c_{ij} > 0$ represent (product) differentiation prices and reflect agent i 's preferences for energy trading. Denote d_i as agent i 's demand and Δg_i as agent i 's renewable energy generation (wind, solar, etc.) which we assume to be a random variable with a CDF, $F_r \in [0, \infty)$. In this work, we do not account for a correlation between agents' random variables while it constitutes an important step for a further research. Then, each agent has

to make a trading decision in the first stage (day-ahead market) about acquiring ($q_i^{da,b}$) or selling ($q_i^{da,s}$) energy at prices $p^{da,b}$, $p^{da,s}$ respectively. At the second stage (real-time market), agents settle imbalances after observing the realization of Δg_i for the prices $p^{rt,b}$ (buying) and $p^{rt,s}$ (selling).

$$\min_{q_i^{da}, q_i^{rt}, \mathbf{q}_i} \overbrace{p^{da,b} q_i^{da,b} - p^{da,s} q_i^{da,s} + \sum_{j \in \Gamma_i} c_{ij} q_{ij}}^{1^{st} \text{ stage costs}} \quad (1a)$$

$$+ \mathbb{E} \left[\overbrace{p^{rt,b} q_i^{rt,b} - p^{rt,s} q_i^{rt,s}}^{2^{nd} \text{ stage costs}} \right] \quad (1b)$$

$$\text{s.t. } q_{ij} + q_{ji} = 0, \quad \forall j \in \Gamma_i \quad (1b)$$

$$d_i = \Delta g_i + \sum_{j \in \Gamma_i} q_{ij} + q_i^{da,b} - q_i^{da,s} + q_i^{rt,b} - q_i^{rt,s} \quad (1c)$$

$$q_{ij} \leq \kappa_{ij}, \quad \forall j \in \Gamma_i \quad (1d)$$

$$q_i^{da,b} \geq 0, q_i^{da,s} \geq 0, q_i^{rt,b} \geq 0, q_i^{rt,s} \geq 0 \quad (1e)$$

Remark 1: For the sake of simplicity, demand d_i is assumed to be deterministic, as d_i is prosumer i 's own demand, and hence may be controlled by i . This reflects the practical case where prosumers can have relatively precise control over their demand (e.g. using flexible demand and/or storage in an energy management system). In the more general case, demand can also be a random variable, as it is subject to factors which can be out of the control of the prosumer (e.g., weather conditions, occupancy, etc). Our results can be extended to this case by considering the forecast of the net load $d_i - \Delta g_i$ (instead of simply Δg_i), which introduces an additional term in the proof of Theorem 1 (net load random variables can have negative values).

Note, that the expectation of the second stage costs is taken with respect to a Gaussian distribution with CDF, F_r , which represents a real distribution of Δg_i . Nevertheless, without full knowledge of this distribution, agent i has access to a forecast (belief) with CDF F_i (Gaussian distribution with the parameters different of those of F_r) about the distribution of Δg_i , which she uses for computing the solution of the problem (1). In the next section, we provide a description of the forecast market, which allows agent i to update her belief to be subsequently used in (1).

C. Forecast Market

We adopt a model for the forecast market from [4], [14] illustrated in Figure 1. In this section we provide a brief high-level description of the properties of the forecast market that are relevant to our model. For the detailed description, we refer to [4]. Let prosumer i be a buyer of the forecast who is interested in improving a forecast (e.g., forecasting algorithm, weather forecast or a generation forecast for their renewable energy asset). For this purpose, the prosumer enters the market by posting a forecasting task for a realization of Δg_i , their own forecast report F_i as a reference for improvement. Each seller of the forecast $j \in I$ reports their forecast f_j along with a wager

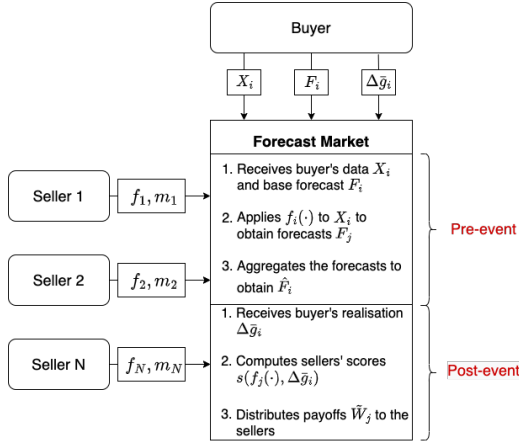


Fig. 1. Proposed data market scheme

$m_j > 0$ defined by themselves, which expresses their confidence on their forecast. We note that the client is also allowed to enter the market as a player with their own forecast report and wager. The client can compete for a relative forecasting skill reward and also influence the resulting forecast. Finally, the forecast market operator aggregates all the forecasts provided by the players, considering their wagers, and delivers the resulting report $\hat{F}_i(\mathbf{m}, \mathbf{f})$ to the client.

After the occurrence of the event, i.e., the time interval for which the forecast is being elicited, the market operator observes the true outcome $\Delta \hat{g}_i$ and evaluates the score of each seller $j \in I$, which shows how “good” was the forecast reported by seller j . Furthermore, the operator also evaluates the utility allocated by the client for the forecast improvement in monetary terms, which then has to be distributed among the players that have contributed to the improvement.

As authors show in [4], this forecast market mechanism enjoys some desirable properties such as budget-balance, anonymity, sybilproofness, truthfulness for the client and an individually rational (IR) for the forecast sellers, adapted from [14]. Note that in our study we investigate IR of the buyers of the forecasts, which is different from the latter. We refer to [4], [14] for more details.

III. COUPLED MARKETS ANALYSIS

In this section we first derive the optimal procurement quantities for the agents, by solving problem (1) as a variant of stochastic inventory management problem complicated by the peer-to-peer trading. Then we focus on two important properties of our connection between the forecast and peer-to-peer market: efficiency and individual rationality. We show conditions for both properties to hold and discuss the limitations of these assumptions.

Denote the residual after first-stage decisions as $r_i := d_i - q_i^{da,b} + q_i^{da,s} - \sum_{j \in \Gamma_i} q_{ij}$ and note that it is non-negative. We first show the following important result about the optimal procurement strategy in the presence of day-ahead and real-time contracts in a market with random renewable generation:

Theorem 1: The residual r_i of agent i after the day-ahead market is given by

$$\begin{aligned} q_i^{da,b} - q_i^{da,s} + \sum_{j \in \Gamma_i} q_{ij} &= d_i - F_i^{-1} \left(\frac{p^{da,b} - p^{rt,s} - \mu_i^{da,b}}{p^{rt,b} - p^{rt,s}} \right) \\ &= d_i - F_i^{-1} \left(\frac{p^{da,s} - p^{rt,s} + \mu_i^{da,s}}{p^{rt,b} - p^{rt,s}} \right) \\ &= d_i - F_i^{-1} \left(\frac{c_{ij} - p^{rt,s} + \zeta_{ij} + \xi_{ij}}{p^{rt,b} - p^{rt,s}} \right). \end{aligned} \quad (2)$$

More precisely,

$$r_i = F_i^{-1} \left(\frac{p^{da,b} - p^{rt,s}}{p^{rt,b} - p^{rt,s}} \right) \text{ or } r_i = F_i^{-1} \left(\frac{p^{da,s} - p^{rt,s}}{p^{rt,b} - p^{rt,s}} \right) \quad (3)$$

if agent i purchases (or sells) electricity on the first stage

Proof: Sketch of the proof: The proof follows classical derivations of first-order stationary conditions for stochastic inventory models. We expand second-stage costs using (1c) to substitute the second-stage variables in the second-stage costs. We then write the KKT conditions, from which we obtain the values for the r_i . \square

The result above expresses the agents’ decision on the day-ahead market in terms of residuals r_i , i.e. the quantities representing the additional purchases that each agent needs to make to balance the uncertainty of the supply after observing the realization of the renewable generation Δg_i . While not providing us the explicit expressions for the decision variables $q_i^{da,b}$, $q_i^{s,da}$, q_{ij} , it allows us to obtain important results about electricity market properties.

A. Efficiency of the Peer-to-Peer Market

When designing market rules, it is important to choose an equilibrium with desirable properties from a set of equilibria (possibly infinite). In our analysis we rely on a notion of Generalized Nash Equilibria and its refinement, Variational Equilibria (VE) as discussed in [13].

Definition 1: A Generalized Nash Equilibrium (GNE) of the game defined by the problems (1) with coupling constraints, is a vector $x_i := (q_i^{da}, q_i^{rt}, \mathbf{q}_i)$ that solves the agents’ problems in (1) or, equivalently, a vector $x_i := (q_i^{da}, q_i^{rt}, \mathbf{q}_i)$ such that $x_i := (q_i^{da}, q_i^{rt}, \mathbf{q}_i)$ solve the system KKT_i for each i .

Definition 2: A Variational Equilibrium (VE) of the game defined by the maximization problems (1) with coupling constraints, is a GNE of this game such that, in addition, the Lagrangian multipliers ζ_{ij} associated to the coupling constraints $q_{ij} + q_{ji} = 0$ are equal, i.e.:

$$\zeta_{ij} = \zeta_{ji}, \quad \forall i \in \mathcal{N}, \forall j \in \Gamma_i \quad (4)$$

By duality theory, ζ_{ij} for $i \in \mathcal{N}, \forall j \in \Gamma_i$ can be interpreted as bilateral energy trading prices [1]. In general, ζ_{ij} might not be aligned with ζ_{ji} , thus leading to non-symmetric energy trading prices between couple of agents. Relying on VE as a solution concept enforces a natural symmetry in the bilateral energy price valuations [1]. To give a condition on VE existence, we first derive the following lemma:

Lemma 1: For any pair of agents (i, j) such that $j \in \Gamma_i$, dual variables ξ_{ij} associated with capacity constraints $q_{ij} \leq \kappa_{ij}$ are equal to zero, $\xi_{ij} = \xi_{ji} = 0$.

Proof: From complementarity conditions for (1), we have that $(q_{ij} - \kappa_{ij})\xi_{ij} = 0$ and $(q_{ji} - \kappa_{ji})\xi_{ji} = 0$. Using the fact that $q_{ij} + q_{ji} = 0$ and the fact that $\kappa_{ij} = \kappa_{ji}$, we rewrite the second equality as $(q_{ij} + \kappa_{ij})\xi_{ji} = 0$. Taking the sum with the first equality, we obtain that $q_{ij}(\xi_{ij} + \xi_{ji}) - \kappa_{ij}(\xi_{ij} - \xi_{ji}) = 0$. Next we note that $\xi_{ij} \neq 0$ only when $q_{ij} = \kappa_{ij}$, thus, $\kappa_{ij}(2\xi_{ji}) = 0$, from which it follows that $\xi_{ji} = 0$. Symmetric reasoning for q_{ji} concludes the proof. \square

Lemma 1 allows us to use KKT conditions for (1) to show the conditions of VE existence. Recap that $r_i := d_i - q_i^{da,b} + q_i^{da,s} - \sum_{j \in \Gamma_i} q_{ij}$ and assume that agent i buys energy on the day-ahead market and agent j sells it on the day-ahead market (the opposite case is considered similarly). Then, at the Variational Equilibrium $\zeta_{ij} = \zeta_{ji}$, thus

$$c_{ij} - c_{ji} = (p^{rt,b} - p^{rt,s}) [F_i(r_i) - F_j(r_j)] = p^{da,b} - p^{da,s},$$

which might raise a question, whether VE exists only if the condition $c_{ij} - c_{ji} = p^{da,b} - p^{da,s}$ holds. It is indeed the case as we show in the next proposition:

Proposition 1: VE of the problem (1) exists iff $c_{ij} - c_{ji} = p^{da,b} - p^{da,s}$ when agent i (j) purchases (sells) electricity on the day-ahead market and if $c_{ij} - c_{ji} = p^{da,s} - p^{da,b}$ in the opposite case.

Proof: Following Definition 1.2 in [13], Variational Equilibrium exists, if there exists $x \in \mathbb{C}$, where $x := (x_1, \dots, x_N)$ - tuple of all the decision variables of all agents and \mathbb{C} - shared constraint set, such that x solve the following variational inequality:

$$F(x)^T(y - x) \geq 0, \quad \forall y \in \mathbb{C}, \quad (5)$$

We assume the existence of x , then, assuming infinitesimal deviations from it we show that $c_{ij} - c_{ji} = p^{da,b} - p^{da,s}$ holds. The other direction follows directly from the KKT conditions. \square

This result gives a necessary condition for VE to exist, which allows us to further prove the efficiency property for our peer-to-peer market coupled with the forecast market. Indeed, solving problem (1) in a centralized way leads to redundancy of the both bilateral trading conditions $q_{ij} + q_{ji} = 0$ and $q_{ji} + q_{ij} = 0$ for a pair of agents (i, j) . Thus, considering VE we can guarantee that the solutions at VE are defined by exactly the same KKT system as the social welfare maximizer, if the distributions used for the solutions of these problems are the same.

First, we prove an interesting result, that shows that there is no impact of the agents' forecasts on the other agents. Note that the final trading costs of agent i consist of two terms: $\sum_{j \in \Gamma_i} c_{ij}q_{ij}$, representing preferences and $\sum_{j \in \Gamma_i} \zeta_{ij}q_{ij}$, representing the bilateral payment for the trade. Then, using KKT conditions, we obtain that

$$\begin{aligned} \Pi_i^q &= \sum_{j \in \Gamma_i} (\zeta_{ij} + c_{ij})q_{ij} = \sum_{j \in \Gamma_i} \left[(p^{rt,b} - p^{rt,s})F_i(r_i) + p^{rt,s} \right] q_{ij} \\ &= \left[(p^{rt,b} - p^{rt,s})F_i(r_i) + p^{rt,s} \right] \sum_{j \in \Gamma_i} q_{ij}, \end{aligned} \quad (6)$$

which, with the fact that $\sum_{j \in \Gamma_i} q_{ij}$ and residual r_i depend solely on agent i 's parameters (expressions (3)) gives us the following result:

Theorem 2: Total cost of agent i depends only on the parameters of agent i . This implies that the forecast market operator can compute utility change of agent i without information from the other agents.

Proof: The proof follows directly from equations (3) and (6). \square

This result is not only interesting by itself, but it is also crucial for the coupling of the forecast market and the peer-to-peer market, as it allows to endogenize the utility of the prosumer introduced by the forecast update without collecting information from the other agents in the system. The seller's scores and payoff distribution then occur as in [4].

Nonetheless, this result does not immediately lead to a market efficiency. Indeed, in order to guarantee it, we have to assume that all the agents report their forecasts to the market operator (not necessarily purchase the forecast).

Theorem 3: If all the agents report their forecasts to the Market Operator (participate in the forecast market), then the VE of (1) coincides with the set of social welfare optima.

Proof: The proof follows from Lemma 1, Proposition 1 and KKT conditions written for peer-to-peer electricity market and centralized formulation of (1). \square

This assumption is supported by the fact that we establish in the next section, more precisely, individual rationality of the coupling between forecast market and peer-to-peer electricity market, which we establish in the next section. We show, that under mild conditions, agents benefit from purchasing the forecasts, thus, they have incentive for participation in the forecast market.

B. Individual Rationality

In this section we show the proof of IR for the case in which the forecasts are given in a form of the truncated Gaussian distribution, defined on \mathbb{R}^+ in order to account for the positive values of Δg_i . As we stated in Remark 1, this analysis can be extended to distribution with support in \mathbb{R} . Computations for the general case are provided in numerical section.

From (2) we obtain that $r_i = F_i^{-1}\left(\frac{\zeta_{ij} + c_{ij} - p^{rt,s}}{p^{rt,b} - p^{rt,s}}\right)$, thus, $\Pi_i^q = (\zeta_{ij} + c_{ij}) \sum_{j \in \Gamma_i} q_{ij}$. From KKT conditions we have that $\zeta_{ij} + c_{ij} = c_i$ for each $j \in \Gamma_i$, where c_i is some constant specific for each agent with $p^{da,s} \leq c_i \leq p^{da,b}$. Then, from (2) we obtain that either c_i is equal to $p^{da,b}$, if agent i buys energy from the backup retailer, or to $p^{da,s}$ otherwise. It allows us to finally write expressions for the total cost imposed on the agent i . First, consider the case when i buys energy from backup retailer on the day-ahead market

$$\begin{aligned} \Pi_i^{total} &= p^{da,b} q_i^{da,b} + p^{da,b} \sum_{j \in \Gamma_i} q_{ij} + \Pi_i^{second} \\ &= p^{da,b} \left[d_i - F_i^{-1}\left(\frac{p^{da,b} - p^{rt,s}}{p^{rt,b} - p^{rt,s}}\right) \right] + \Pi_i^{second}, \end{aligned} \quad (7)$$

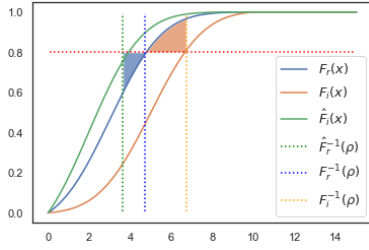


Fig. 2. Illustration for condition (10). Blue area should be smaller than the orange one.

where Π_i^{second} is given by

$$\Pi_i^{second} = \overbrace{p^{rt,b} r_i F_r(r_i) + p^{rt,s} r_i (1 - F_r(r_i))}^1 - \overbrace{p^{rt,b} \int_0^{r_i} \Delta g_i f_r(\Delta g_i) d\Delta g_i - p^{rt,s} \int_{r_i}^{\infty} \Delta g_i f_r(\Delta g_i) d\Delta g_i}^2$$

where F_r (f_r) denotes CDF (PDF) of a real distribution of Δg_i . It means that Π_i^{total} gives an expected cost of agent i who takes r_i as a first stage decision (r_i denotes residual after the first stage). Breaking down the second stage costs and summing it all together:

$$\begin{aligned} \Pi_i^{total} &= p^{da,b} [d_i - r_i] - (p^{rt,b} - p^{rt,s}) F_r(r_i) \\ &+ (p^{rt,b} - p^{rt,s}) \mathbb{E}_r[\Delta g_i | \Delta g_i \geq r_i] (1 - F_r(r_i)) \\ &+ p^{rt,s} r_i - p^{rt,b} \mathbb{E}_r[\Delta g_i] \end{aligned} \quad (8)$$

which gives us expected cost of an agent i .

Expression (8) allows us to consider the effects of the forecast market on the peer-to-peer market in an expectation with respect to the real distribution of Δg_i . We would like to show that

$$\mathbb{E}_r[\hat{\Pi}_i^{total}] \leq \mathbb{E}_r[\Pi_i^{total}], \quad (9)$$

if distribution \hat{F}_i is "better" than F_i . Intuitively, for one shot game it should compare two distributions by the amount of probability mass concentrated around the realization of a random variable Δg_i .

Consider an agent i who has an initial forecast about the distribution of Δg_i with CDF F_i of a Gaussian distribution and a 'better' forecast with CDF \hat{F}_i . Now, fix prices $p^{rt,b} > p^{da,b} > p^{da,s} > p^{rt,s}$ and denote $\rho := \frac{p^{da,b} - p^{rt,s}}{p^{rt,b} - p^{rt,s}}$. Assume that agent i buys energy on the day-ahead market if she uses F_i or \hat{F}_i or F_r (this can be expressed as $F_i^{-1}(\rho), \hat{F}_i^{-1}(\rho), F_r^{-1}(\rho) \leq d_i$). Situation in which i sells energy on the day-ahead market is considered similarly. Moreover, denote $\hat{r}_i := \hat{F}_i^{-1}(\rho)$. Then, for the Gaussian case, we write $r_i = \mu_i + \sigma_i \Phi(\rho)$ and $\hat{r}_i = \hat{\mu}_i + \hat{\sigma}_i \Phi(\rho)$. Now, plugging (8) into (9), using the fact that $\mathbb{E}_r[\Delta g_i | \Delta g_i \geq r_i] (1 - F_r(r_i)) = \int_{r_i}^{\infty} x f_r(x) dx$, where $f_r(x)$ is the PDF of a real distribution of Δg_i , and integrating by parts, we can write the following condition on the distributions:

$$\int_{\mu_i + \sigma_i \Phi(\rho)}^{\mu_r + \sigma_r \Phi(\rho)} (\rho - F_r(x)) dx \geq \int_{\hat{\mu}_i + \hat{\sigma}_i \Phi(\rho)}^{\mu_r + \sigma_r \Phi(\rho)} (\rho - F_r(x)) dx \quad (10)$$

First, note that $\rho = F_r(F_r^{-1}(\rho))$, thus, inequality clearly holds when $F_i^{-1}(\rho) \leq \hat{F}_i^{-1}(\rho) \leq F_r^{-1}(\rho)$ or when $F_i^{-1}(\rho) \geq$

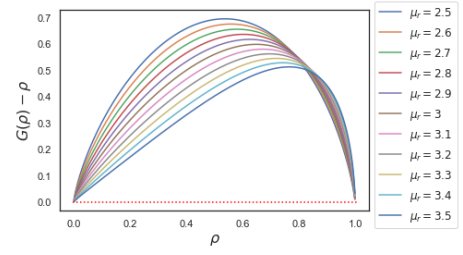


Fig. 3. $G(\rho) - \rho$ for different distances between distributions. Individual Rationality holds \forall values of ρ .

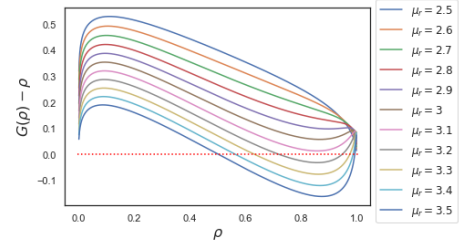


Fig. 4. $G(\rho) - \rho$ for different distances between distributions with smaller variance.

$\hat{F}_i^{-1}(\rho) \geq F_r^{-1}(\rho)$. We next assume that $\hat{F}_i^{-1}(\rho) \leq F_r^{-1}(\rho) \leq F_i^{-1}(\rho)$, while the opposite case can be considered similarly. From (10) we directly obtain that

$$\rho \leq \frac{1}{\mu_i - \hat{\mu}_i} \int_{\mu_i + \sigma_i \Phi(\rho)}^{\hat{\mu}_i + \hat{\sigma}_i \Phi(\rho)} \quad (11)$$

which defines the relationship between the quantiles of the forecasts and CDF of a real distribution of Δg_i . For the sake of illustration, we consider shifted Gaussian distributions with the same variance σ , in which case condition (11) reduces to

$$\begin{aligned} \rho \leq G(\rho) &:= \frac{\mu_i - \mu_r + \sigma \Phi^{-1}(\rho)}{\mu_i - \hat{\mu}_i} \Phi\left(\frac{\mu_i - \mu_r}{\sigma} + \Phi^{-1}(\rho)\right) \\ &- \frac{\hat{\mu}_i - \mu_r + \sigma \Phi^{-1}(\rho)}{\mu_i - \hat{\mu}_i} \Phi\left(\frac{\hat{\mu}_i - \mu_r}{\sigma} + \Phi^{-1}(\rho)\right) \\ &+ \sigma \left(\phi\left(\frac{\mu_i - \mu_r}{\sigma} + \Phi^{-1}(\rho)\right) + \phi\left(\frac{\hat{\mu}_i - \mu_r}{\sigma} + \Phi^{-1}(\rho)\right) \right) \end{aligned}$$

which can be easily evaluated numerically. Figure 3 demonstrates the values of $G(\rho) - \rho$ for different μ_r while μ_i and $\hat{\mu}_i$ are fixed and are equal to 5 and 2 respectively and $\sigma = 2$. The closer $\hat{\mu}_i$ to μ_r comparing to $\mu_i - \mu_r$, the bigger admissible values of ρ are. Figure 4 demonstrates the values of $G(\rho) - \rho$ for Gaussian distributions with $\sigma = 0.5$. We see that the closer 'better' forecast to the real distribution, the higher the value of ρ for which we can guarantee Individual Rationality.

IV. NUMERICAL EXPERIMENTS

In this section, we provide several numerical examples to provide an illustration on the proposed market model coupling and to numerically demonstrate the properties of the proposed model on the real-world data example. The code for the experiments and the values of all the parameters are provided in [15].

a) *Experimental setup*: We use residential data provided by Pecan Street [16] for Austin, Texas. The data consists of 15-minutes intervals specifying renewable generation, load and facilities energy consumption for 25 individual homes. We sample the distribution of scenarios for RES-based generation and demand from the generation data and aggregated consumption respectively. The electricity market prices are chosen arbitrarily to be $p^{da,b} = 0.3$ [\$/kWh], $p^{da,s} = 0.2$ [\$/kWh], $p^{rt,b} = 0.4$ [\$/kWh] and $p^{rt,s} = 0.1$ [\$/kWh]. These prices were selected for demonstration purposes only, and they can be set to any value as long as they maintain the assumed order $p^{rt,b} > p^{da,b} > p^{da,s} > p^{rt,s}$.

Each agent on the peer-to-peer electricity market can arrive at the wagering based forecasting market, described in Section II-C ([4] for detailed description), as a client. The client submits the task of forecasting the next 15-minutes of renewable energy generation Δg_i . As a historical data X_i , agent i provides observations of RES-generation [15], from which the sellers build distributions, subsequently used in a forecast \hat{F}_i . It follows that the forecast sellers utilize same data but different models (forecasting skill) to construct the forecasts. We assume that there are 25 sellers in the forecast selling market.

To obtain an aggregated forecast \hat{F}_i we use quantile averaging (QA) described in [4]. In order to showcase the differences in the forecast only, equal weight values of $m_n = 1$ are assigned for all n in $\{1, \dots, 25\}$. The methodology for selecting wagers and their effect are thoroughly discussed in [4]. After the task and the forecasts are submitted to the Data Trading Platform, the forecast market operator evaluates the scores of submitted forecasts compensates accordingly.

Each buyer of the forecast (an agent in the peer-to-peer market) is equipped with a base forecast in a form of truncated normal distribution with a mean equal to a sample mean, computed from historical observations. Variance is computed using the scaling factor of 2 and lies in $[1.455, 2.452]$.

We test two types of the forecasts available to the sellers on the forecast market, corresponding to two settings: (i) In the first setting, we employ truncated normal distribution with the variance smaller than those in the base forecasts of the buyers (prosumers). The means remain the same, while the variance for the 'updated' forecasts now lies in a range $[0.364, 0.852]$. (ii) In the second setting, we test empirical distributions built from historical data as the 'real' forecasts: each seller has access to a distribution $i \in \{1, \dots, 25\}$, from which we then draw Δg_i as a realization of a random variable, corresponding to RES-based generation of agent i .

Finally, we examine a scenario where only one (or two) prosumers purchase the forecast on the forecast market. Due to Theorem 2 we do it without any loss of generality, as the results can be easily extended to any number of buyers. This allows us to concentrate on the impact of the forecast market and the individual rationality property on each agent's level.

b) *Results*: We first illustrate the conditions for IR on the real data. Figure 5 illustrates the distributions (forecasts) used for agent 1. (i) 'Aggregated' - aggregated forecast \hat{F}_i using empirical distributions, (ii) 'Real' - empirical distribution of Δg_i , (iii) 'Base' - truncated normal distribution, (iv) 'Updated' -

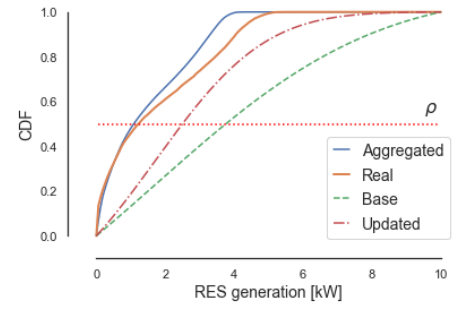


Fig. 5. CDF of different distributions (forecasts) of agent 1.

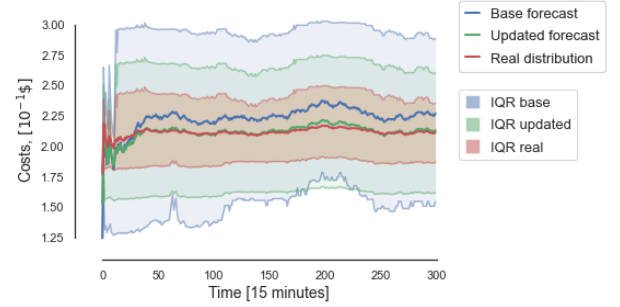


Fig. 6. Sliding mean costs of agent 1 in electricity market with different distributions used to estimate expected costs. Interquartile range (IQR) for each distribution is represented by the shaded area on the graph.

truncated normal distribution with reduced variance. Note that the distance between quantiles of these distributions and the empirical distribution calculated at $\rho = \frac{p^{da,b} - p^{rt,s}}{p^{rt,b} - p^{rt,s}}$ is the lowest for the aggregated forecast, while the distance between the updated forecast is lower than that for the base one.

To represent the costs associated with different distributions used as the forecast of an agent, we use a sliding mean. Figure 6 illustrates the costs of agent 1 on the electricity market associated with different distributions used by the agent to compute the values of her decision variables. We can observe that while the costs associated with a base forecast are the largest, the costs associated with a 'real' forecast and 'updated' one are very close, with the latter rarely exceeding the former. We also observe that the interquartile range (IQR) associated with the distributions decreases as the distribution is closer (in a sense of condition (10), as illustrated on Figure 5) to the real one. It means that even if in some time intervals the 'best' forecast might result in higher costs for the agent, the volatility in the costs of the agent is reduced, which indicates reduced volatility in the electricity market.

One can intuitively guess that a better forecast does not guarantee individual rationality in the one-shot game and the purchase of the forecast on a forecast market can even increase the costs of the agent in the short term. Figure 7 demonstrates that even the 'best' forecast leads to the increase in the costs of agent 1 in the beginning, within ~ 50 of 15-minute intervals. Nevertheless, the total effect of the forecast purchase affects the outcome for the agents positively in a long-term horizon. We observed that it decreased real costs for the agents purchasing

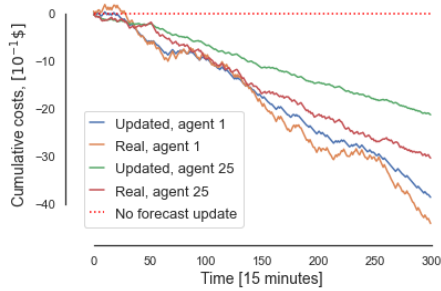


Fig. 7. Cumulative changes in the costs of agents 1 and 25 due to the improvements in their forecasts.

the forecast, while the costs of the other agents remained the same, as illustrated by the red dotted line in Figure 7. Note that the cumulative costs associated with the 'real' forecasts are lower than those associated with the 'updated' one, which, with Figure 5 illustrates the effect of the 'closeness' of the distributions which we discuss in Section III-B.

We illustrate the profits (losses) of the sellers on the forecast market entailed by the forecast purchase by agent 1 in Figure 8. For the sake of presentation, we illustrate only the profits (losses) of the top (bottom) two scoring sellers in the forecast market. We measure the utility obtained by the agent 1 from purchasing the forecast by measuring the difference between the agent's cost associated with a 'base' forecast and the 'updated' one in this case. In our proposed coupling mechanism, this is carried out (as well as scoring allocation for the forecast providers) by the forecast market operator. On average, sellers 11 and 21 performed better than the others, and the opposite holds for sellers 3 and 14. Still, as is shown by the bar plot, in some instances these agents suffer some losses, caused by misprediction.

V. CONCLUSION

In this work, we presented a framework for coupling electricity peer-to-peer and forecast markets. We addressed the main research questions arising from this coupling: (i) how to evaluate the impact of the forecast update on the agent's utility, (ii) how it improves the operation of the electricity market, and (iii) started the investigation of what are the incentives for the agents in the electricity market to purchase the forecasts. We proved that i) the impact of the forecast purchase on the prosumers' utility can be computed without information from the other prosumers; ii) if all the prosumers participate in the forecast market, then the electricity market economic efficiency is achieved; and iii) for the agents there exist incentives to participate in the forecast market, if the forecasts are given in a form of the Gaussian distribution, i.e., the purchase of the forecast leads to decreased costs, given condition on the 'distance' between the distributions. Thus, the importance of this result lies not only in the fact that it is profitable for the agent to purchase a forecast but also in that it allows for achieving electricity market efficiency.

This paper serves as a starting point in exploring the connection between electricity and forecast markets, highlighting

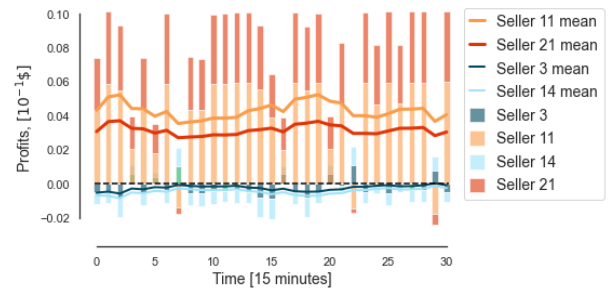


Fig. 8. Profits of the top two and bottom two scoring sellers in the forecast market. Bars represent obtained profit (loss) in each 15-minute interval, while sliding means are depicted as the lines of a corresponding color.

various key dimensions which can further improve the efficiency of this coupling. The first direction for further research is to extend the presented results to the general case, beyond Gaussian distributions. Another direction would focus on determining a suitable metric for comparing probabilistic forecasts that ensures individual rationality in the electricity market.

REFERENCES

- [1] H. Le Cadre, P. Jacquot, C. Wan, and C. Alasseur, "Peer-to-peer electricity markets: from variational to generalized nash equilibrium," *European Journal of Operational Research*, vol. 282, no. 2, pp. 753–771, 2020.
- [2] F. Moret, P. Pinson, and A. Papakonstantinou, "Heterogeneous risk preferences in community-based electricity markets," *European Journal of Operational Research*, vol. 287, no. 1, pp. 36–48, 2020.
- [3] J. Nair, S. Adlakha, and A. Wierman, "Energy procurement strategies in the presence of intermittent sources," *ACM SIGMETRICS Performance Evaluation Review*, vol. 42, 2014.
- [4] A. A. Raja, P. Pinson, J. Kazempour, and S. Grammatico, "A market for trading forecasts: A wagering mechanism," *International Journal of Forecasting*, 2023. Accepted/In press.
- [5] F. Petropoulos, D. Apiletti, V. Assimakopoulos, M. Z. Babai, D. K. Barrow, S. Taieb, C. Bergmeir, R. Bessa, J. Bijak, *et al.*, "Forecasting: Theory and practice," *International Journal of Forecasting*, vol. 38, pp. 705–871, 2022.
- [6] G. Perakis and G. Roels, "Regret in the newsvendor model with partial information," *Operations Research*, vol. 56, pp. 188–203, 2008.
- [7] A. Agarwal, M. Dahleh, and T. Sarkar, "A marketplace for data: An algorithmic solution," in *Proceedings of the 2019 ACM Conference on Economics and Computation*, pp. 701–726, ACM, 2019.
- [8] J. Wolfers and E. Zitzewitz, "Prediction markets," *Journal of Economic Perspectives*, vol. 18, no. 2, pp. 107–126, 2004.
- [9] J. W. Messner and P. Pinson, "Online adaptive lasso estimation in vector autoregressive models for high dimensional wind power forecasting," *International Journal of Forecasting*, vol. 35, pp. 1485–1498, 2019.
- [10] T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, and R. J. Hyndman, "Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond," *International Journal of Forecasting*, vol. 32, no. 3, pp. 896–913, 2016.
- [11] H. Fan, X. Zhang, S. Mei, *et al.*, "M2gsnet: Multi-modal multi-task graph spatiotemporal network for ultra-short-term wind farm cluster power prediction," *Applied Sciences*, vol. 10, no. 21, p. 7915, 2020.
- [12] P. T. Harker, "Generalized nash games and quasi-variational inequalities," *European Journal of Operational Research*, vol. 54, no. 1, pp. 81–94, 1991.
- [13] A. A. Kulkarni and U. V. Shanbhag, "On the variational equilibrium as a refinement of the generalized nash equilibrium," *Automatica*, vol. 48, no. 1, pp. 45–55, 2012.
- [14] N. S. Lambert, J. Langford, J. Wortman, Y. Chen, D. Reeves, Y. Shoham, and D. M. Pennock, "Self-financed wagering mechanisms for forecasting," in *Proceedings of the 9th ACM Conference on Electronic Commerce*, pp. 170–179, ACM, 2008.
- [15] I. Shilov, "Data market repository." https://github.com/ishilov/data_market, 2023.
- [16] "Pecan street dataport." <https://dataport.pecanstreet.org/>, 2022.