

Agent-based Decentralized Optimal Charging Strategy for Plug-in Electric Vehicles

Milad Latifi, Amir Rastegarnia, Azam Khalili, and Saeid Sanei, *Senior Member, IEEE*

Abstract—This paper presents a game theoretic decentralized electric vehicle charging schedule for minimizing the customers' payments, maximizing the grid efficiency, and providing maximum potential capacity for ancillary services. Most of the available methods for electric vehicle charging assume that the customers are rational, there is low-latency perfect two-way communication infrastructure without communication/computation limitation between the distribution company and all the customers, and they have perfect knowledge about the system parameters. To avoid these strong assumptions and preserve the customers' privacy, we take advantages of the regret matching and the Nash Folk theorems. In the considered game, the players (customers) interact and communicate locally with only their neighbors. We propose a mechanism for this game which results in a full Nash Folk theorem. We demonstrate and prove that the on-off charging strategy provides maximum regulation capacity. However, our mechanism is quite general, takes into account the battery characteristics and degradation costs of the vehicles, provides a real time dynamic pricing model, and supports the vehicle-to-grid (V2G) and modulated charging protocols. Moreover, the developed mechanism is robust to the data disruptions and takes into account the long/short term uncertainties.

Index Terms—Plug-in electric vehicles (PEVs), decentralized charging, Nash Folk strategy, regret matching.

I. INTRODUCTION

PLUG-IN electric vehicles (PEVs) will soon electrify the entire structure of the future transportation systems [1]. Reducing dependency on fossil fuels, reducing emissions of greenhouse gases, energy and cost-saving, and better utilizing of renewable energy sources are among the leading reasons that PEVs are increasing in popularity. However, uncoordinated charging of PEVs could have adverse effects on reliability, stability, and efficiency of the power system. These adverse effects include increasing the system peak load or creating new sub-peaks, increasing power losses, decreasing the load factor, and causing system voltage deviation and overloading of transformers [2]. So, it is critical to develop a well-designed charging coordination mechanism to alleviate the undesirable effects and enhance the benefits of this electrification. From this point of view, several studies have been carried out in recent years to provide optimal charging schedules for PEVs [3]–[14].

Manuscript received XX, 2016; revised XX, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was XXXX XXXX. M. Latifi, A. Khalili, A. Rastegarnia are with the Department of Electrical Engineering, Malayer University, Malayer, 65719-95863, Iran (email: miladlatifi98@gmail.com; rastegarnia@malayeru.ac.ir; khalili@malayeru.ac.ir).

S. Sanei is with School of Science and Technology, Nottingham Trent University, Nottingham, NG11 8NS, UK (email: saeid.sanei@ntu.ac.uk). Digital Object Identifier XXXXX/XXXXXXX

A. Related Works

The authors of paper in [3] identified and analyzed the relationships and the mutual influence between the feeder losses, the load factor, and the load variance parameters. They developed optimal charging algorithms for each of the parameters, while minimizing the impact of PEV charging on the connected distribution system. They denoted by the simulation results on the two test systems that these relationships approximately hold independent of system topology. Further, they showed that it is more beneficial in terms of problem convexity and the computational complexity to use load factor or load variance as the objective function rather than system losses. He *et al.* introduced a global convex optimization problem in [4], which aimed to minimize the total cost for charging all the PEVs within the day. As the proposed global optimal solution needs information about the future base loads and the arrival times and the charging periods of the PEVs that will arrive in the future time of the day, it is impractical in real world. So, they assigned a local optimization problem to the PEVs scheduling in each local group, which is resilient to the dynamic PEV arrivals and aims to minimize the total cost of the PEVs in the current ongoing groups in an independent and distributed manner. They demonstrated by the numerical simulations that the locally optimal scheduling scheme can achieve a close performance compared to the globally optimal scheduling scheme.

The work in [5] formulated a game theoretical model to characterize the interactions among the PEVs and the aggregator in a V2G protocol. Using this model, the authors of this paper designed a mechanism to achieve optimal frequency regulation performance in a distributed fashion which benefits both the customers and the power grids. To encourage the PEVs to participate in the frequency regulation service, a dynamic pricing policy is considered in this paper. Moreover, they provided a new model explaining how a backup battery bank can be deployed in an aggregator to maintain a stable regulation capacity. Through analyzing the Nash equilibrium in the vehicle-to-aggregator interaction games, they showed that the proposed decentralized mechanism works as efficiently as the centralized controlled methods. Reference [6] formulated a decentralized iterative algorithm to schedule the PEV charging to fill the valleys of the aggregated load demand curve, which results in a valley-filling charging profile. In each iteration, PEVs update their charging profiles according to the control signal broadcast by the utility company, and the utility company alters the control signal to guide their updates. This procedure makes possible of real-time implementation and tracking a desirable given load profile. The proposed method preserves optimality (i.e., make the aggregate demand as flat

as it can possibly be) even when the PEVs do not necessarily update their charging profiles in every iteration. The proposed algorithm in this work only requires each PEV solving its local problem, hence its implementation requires low computation capability.

A distributed algorithm for charging control of PEVs was proposed in [7] based on the regret minimization, while the utility functions of various PEV owners are not identical, or precisely known to the distribution company. This algorithm requires only one-way communication from the distribution company to the customers, does not require infrastructure that can support low latency two-way communication, and there is no need for the customers to send their private data to the distribution company. However, the charging profiles in this work are sought to be optimized over the set of profiles that do not vary from day to day, while the algorithm convergence to the optimal charging schedule is slow. Based on the non-cooperative game theory, a new price-driven charging control scheme was developed in [8] to coordinate large scale PEVs without compromising the security of the distribution network. The aim of this work is to minimize the cost of individual PEV owners, considering the overload constraints in the distribution feeders which is tested on the IEEE 13-bus system. A Newton-type fixed point method was formulated in this work to find a better Nash equilibrium of the game model at a superlinear convergence rate. Furthermore, an accelerated gradient method was proposed to tackle the sub-problem for each customer's best response which is implemented in a distributed way in order to protect customer's privacy.

The impact of the PEVs demand on the system electricity price was studied in [9]. The optimal scheduling of the individual PEV controller considering the actions of other PEVs in the game is developed with the PEV driving pattern distribution. An aggregative game model was proposed by this paper for modeling interaction between the PEVs during the day-ahead charging management. An optimization method is developed to calculate the equilibrium of the game model through quadratic programming and the existence and uniqueness of its pure strategy Nash equilibrium was proved using Karush-Kuhn-Tucker (KKT) optimality conditions. Case studies with the proposed game model were carried out using real world driving data from the Danish National Travel Surveys. The impacts of the PEV driving patterns and price forecasts on the EV demand with the proposed game model were also analyzed in this work. In [10], Liu *et al.* provided an online and near-optimal dynamic stochastic linear programming scheme to optimize electric vehicle recharging costs and increase the reliability of system taking into account load, electricity pricing, and renewable energy generation uncertainties. At first offline stage of this work, a day-ahead problem is solved. Subsequently, at the second online stage, the proposed framework uses offline solutions, collects real-time system data, and the stochastic market parameters, and adjusts recharging schedules to obtain a better recharging scheme once system uncertainties are revealed. Since the aggregator communicates to each PEV only once, their method needs a low bandwidth communication infrastructure and can be used in a wide area. Further, the proposed method is robust to variations in different stochastic parameters and benefits both users and the power

utility.

In another research [11], an optimal PEV charging schedule has been formulated assuming that the deriving pattern and future charging demand of the PEVs are not known a priori, but their statistical information can be estimated. The cost of PEV charging was defined as a general strictly convex increasing function of the instantaneous load demand, and was demonstrated that minimizing such a cost leads to a flattened load demand. This problem was formulated as a finite-horizon dynamic programming and a model predictive control (MPC) was provided to avoid the prohibitively high complexity of solving such a problem. The computational complexity and performance gap between the near optimal solution of the MPC-based approach and the optimal solution for any distributions of exogenous random variables were rigorously analyzed in this paper. An optimal charging problem of the PEVs on demand side management (DSM) was formulated in [13] to meet the power system interests, such as reducing the generation supply cost, while respecting each PEV's charging constraint. A distributed consensus initialization-free algorithm was used in this work to estimate the mismatch between all the PEVs allocated powers and the total available charging power on the directed graph. With these mismatch estimation, a non-smooth analysis-based dynamic system was adopted by the authors to dynamically update the PEVs charging power. The proposed multi-agent system strategy is robust to the time-varying available charging power, plug-and-play operations, and single-link failures. Wang *et al.* in [14], introduced a hybrid two-stage centralized-decentralized charging control scheme to reduce the energy cost and guarantee the system stability. On the centralized stage, an offline optimal scheduling approach was presented aiming at minimizing the energy cost while satisfying the charging requirements of the PEVs. To deal with the system dynamics and uncertainties, a real time MPC-based adaptive scheduling strategy is developed by the authors to determine the near optimal PEV charging profiles. Then, on the decentralized stage, the interactions between the PEVs and the charging system controller is modeled as a leader-follower non-cooperative Stackelberg game in which the system controller acts as the leader and the PEVs act as followers. It was shown in this work that by adopting the proposed decentralized charging algorithm, the communication burden between PEVs and the system controller is low and the charging scheme is robust to poor communication channels.

B. Contributions

The main contributions of this paper are summarized as follows.

Designing a novel fully decentralized game theoretic model: Due to the interdependence between PEVs' actions (a PEV action affects other PEVs' payoffs), a game theoretic model for the PEVs charging is designed guarantying the local optimality for each PEV and the global optimality for the system aggregator. As the centralized coordination of large fleet of PEVs is costly, not secure, not reliable, and complex process, a fully decentralized scalable approach is proposed for the first time and its convergence to the Nash Folk equilibrium is proved. In the developed framework, the interactions between PEVs are done locally with neighbors using

an arbitrary-private topology. So, the privacy of the customers are preserved, there is not any vulnerable central point, and the framework is robust to link/node failures. Further, unlike the game theoretic solutions in the literature which do not consider the adversary customers and malfunction data, any customer action deviation or cyber-attack can be identified in the proposed framework. In the mean time, as the emergence of PEVs to the system affects the electricity market price, we formulate a dynamic real time pricing policy in which the customers take into account their impacts on the price.

Improving long/short term global performance: Due to the several uncertainty sources (e.g., unexpected plug-in/out the PEVs, base load, and baseline price) the day-ahead scheduling solution may get away from the optimal point. So, we adjoin the regret matching strategy (to improve the long term performance) and the MPC strategy (to improve the short term performance) with our framework. We also use Bayesian inference and conditional value at risk techniques to tackle the uncertainties in driving behaviors of the customers. Beside considering the customers' payments and battery degradation costs, the proposed method results in a valley-filling charging profile, minimizes the power losses and the load variance, and maximizes the load factor and total regulation capacity.

Supporting universal charging protocols: Charging the PEVs with faster rate and lower time by drawing the maximum charging power from the charging pole [12], prolonging the PEV battery's lifetime with constant power feeding [15], and requiring smaller communication overheads to contact (switch off or on) with a small subset of PEVs [16]. Further, with this method the infrastructures only need to work with one charging rate which needs very cheap infrastructure and services. In addition to the desirable effects of on-off charging strategy, rather than modulating the charging rate, we prove that the on-off strategy results in providing the maximum regulation capacity. However, our proposed method supports the modulated and V2G charging protocols too.

The proposed real-time universal charging scheduling implemented in the fully decentralized manner have potential application in the large power systems with too many customers and the micro-grids equipped with the renewable power generation similar to work in [17]. The model considered in this paper is completely general and can be implemented by different residential/commercial/industrial/organizational customers having one or a group of the PEVs with some minor manipulations.

Notation: Throughout the paper $|\cdot|$ denotes the cardinality of the corresponding set, $\Pr\{\cdot\}$ denoted the probability operator, $\mathbb{E}[\cdot]$ denotes the expectation operator, $[a]^+ = \max\{0, a\}$, and $\langle \cdot, \cdot \rangle$ denotes the inner product. For λ -strongly convex function $L(x)$ with respect to a norm $\|\cdot\|$, $\|\cdot\|_*$ denotes the norm that is dual to $\|\cdot\|$. We denote the Bregman divergence with respect to $L(x)$ as $D_L(x, \hat{x}) := L(x) - L(\hat{x}) - \langle \nabla L(\hat{x}), x - \hat{x} \rangle$, and $\nabla L^{-1}(x)$ denotes the inverse mapping of gradient $\nabla L(x)$.

The rest of this paper is organized as follows. The system components is described in Section II. The PEV energy consumption scheduling problem is formulated in Section III. Our proposed method is introduced in Section IV. Section V provides simulation results and we conclude the paper in Section VI.

II. SYSTEM MODEL

In liberalized electricity markets, the retailers buy electricity from the wholesale market and sell it to their customers in a way to maximize social welfare [18]. Besides, in the liberalized ancillary service market, to facilitate PEVs participation and satisfy the minimum regulation capacity (e.g., 0.1 MW), the aggregators are served as an interface between the independent system operator (ISO) and a fleet of PEVs [12]. The more the PEVs can provide regulation capacity, the more they can acquire financial rebate. On the day before the operation hours, each PEV owner determines the amount of regulation capacity which can or is willing to provide. The aggregator gathers all the regulation capacities and submits its bid to the ISO accordingly. In one hand, the retailer must announce the total future load demand of its customers to the wholesale market before starting the scheduling horizon H (e.g., one day with $H = 24$ hours). This load profile is needed by the wholesale market for energy management and power balancing issues. On the other hand, the aggregator has to sign a contract with the ISO based on the expected storage capacity of its associated PEV fleet before starting the scheduling horizon. So, it is necessary for the retailer (aggregator) to determine a proper framework to reduce the customers' payment (increase the PEVs regulation capacity) in a day-ahead manner¹.

A. Constraints

Consider set \mathcal{K} of PEVs belonging to different residential customers subscribing to the same aggregator. By assuming that each customer has one PEV, the number of the PEVs (customers) is denoted with $K = |\mathcal{K}|$. As the most energy consumption schedules are implemented in two stages (first day ahead programing and then real time programing), we divide a scheduling horizon H into several time slots $\mathcal{H} \triangleq [1, 2, \dots, H]$. The PEVs are schedulable only when they are at home and plugged into the power system. So, there is a schedulable window $\mathcal{H}_k \triangleq [\alpha_k, \dots, \beta_k] \subseteq \mathcal{H}$ for each PEV k , where α_k is the plugged in time and β_k is the last time slot after which the PEV leaves the house. Let x_k^h denote the power rate at which PEV of customer $k \in \mathcal{K}$ is charged ($x_k^h > 0$) or discharged ($x_k^h < 0$) at slot $h \in \mathcal{H}$ (i.e., the energy consumption at each slot). Also, the total inelastic (non-PEV) load demand of the customer in slot h is denoted with l_k^h . To coordinate charging process of the PEV's battery of each customer the following constraints must be satisfied:

$$E_k^{h+1} = E_k^h + \eta_k \frac{x_k^h}{B_k^{cap}}, \quad E_k^{min} \leq E_k^h \leq E_k^{max}, \quad \forall h \in \mathcal{H} \quad (1)$$

$$\sum_{h \in \mathcal{H}_k} \eta_k \frac{x_k^h}{B_k^{cap}} = E_k^d, \quad x_k^h = 0, \quad \forall h \notin \mathcal{H}_k \quad (2)$$

where E_k^h is the state of charge (SOC) of PEV k at slot h , $\eta_k \in (0, 1]$ is the energy conversion efficiency of charging/discharging, B_k^{cap} [kWh] is the maximum energy that the battery can store (storage capacity), E_k^{min} is the lower bound

¹Throughout the paper, we assume that both the retailer and aggregator are the same entity and the terms "retailer, aggregator" and "customer, PEV owner" are used interchangeably.

of the battery's SOC, E_k^{max} is the upper bound of the battery's SOC, and E_k^d is the desired SOC at the end of scheduling horizon H . The first term of (1) is the dynamic equation of the battery's SOC, and the second implies that the SOC of each battery is bounded. For example, we can let $E_k^{min} = 0\%$ and $E_k^{max} = 100\%$. That means, the battery is allowed to be depleted in discharging mode and can be charged to the rated capacity. The first term of (2), imposes the scheduling program to provide the required energy level for PEV's traffic at the departure time (end of the scheduling horizon), and the second term denotes that when the PEV is unplugged, it cannot consume power. So, the feasible set for PEV k 's charging schedule is defined as:

$$\mathcal{X}_k := \{\mathbf{x}_k \mid (1) \text{ and } (2)\} \quad (3)$$

where $\mathbf{x}_k \triangleq [x_k^1, \dots, x_k^H]$ and each charging schedule is acceptable if and only of $\mathbf{x}_k \in \mathcal{X}_k$.

B. Price Function

In the proposed strategies in the demand response (DR) literature, the price signals are used as the main tool to shape the load demand curve [19]. To make an optimal decision in these strategies, the customers need to forecast the price parameters before the scheduling horizon. However, most of the strategies assume that the customers are *price-taker* (PT) [20]. That means, the behavior of the active (who participate in the DR programs) customers does not affect the wholesale price signal. However, when there are a considerable number of active customers (e.g., PEVs) in the grid, their power consumption behavior will be comparable to the conventional demand and inevitably influence the spot prices in the day-ahead market. So, in this paper, the PEVs are considered to be *price-participant* (PP), meaning that, they consider their action's impact on the spot price signal coming from the retailer. To do this, the predicted price at slot h is modeled as follows:

$$\mathcal{P}^h = p^h + \lambda^h \sum_{k \in \mathcal{K}} x_k^h \quad (4)$$

where p^h is the foretasted baseline price at slot h due to the inelastic demand and $\lambda^h \sum_{k \in \mathcal{K}} x_k^h$ is incurred price (say spot price) due to the flexible demand with price sensitivity coefficient λ^h [9].

C. Cost Function

The most important purpose of each PEV owner to participate in the charging schedule is reducing his payment as much as possible. However, frequent charging/discharging the PEV's battery to reduce the payment, also reduces the battery's lifetime. Considering appropriate storage and operation cost functions can be interpreted as imposing a soft constraint to prolong the lifetime preventing large variations of the stored energy. As the lithium-ion type of battery, has been widely applied in electrical vehicles [21], we define the following cost function for customer k incorporating the lithium-ion battery degradation cost:

$$f_k(\mathbf{x}_k) = \sum_{h \in \mathcal{H}} \mathcal{P}^h (x_k^h + l_k^h) + \phi_k^s + \phi_k^f \quad (5)$$

where

$$\begin{aligned} \phi_k^s &= \sum_{h \in \mathcal{H}} \left(\gamma_k^{s1} (x_k^h)^2 + \gamma_k^{s2} |x_k^h| + \gamma_k^{s3} \right), \phi_k^f = \sum_{h=2}^H \gamma_k^f (x_k^h - x_k^{h-1})^2, \\ \gamma_k^{s1} &= 10^6 p^{cel} \gamma^4 / (M_k V^{nom}), \gamma_k^{s2} = 10^3 p^{cel} (\gamma^2 + \gamma^6 V^{nom}), \\ \gamma_k^{s3} &= M_k p^{cel} V^{nom} (\gamma^1 + \gamma^3 V^{nom} + \gamma^5 (V^{nom})^2 + \gamma^7 (V^{nom})^3) \end{aligned} \quad (6)$$

where p^{cel} represents the price of a single energy unit of a battery cell, M_k is the number of cell units in PEV k , and V^{nom} is the nominal value of the open circuit voltage of the battery cell unit [22]. The parameters $\gamma^i, i = 1, 2, \dots, 7$ specified as -1.148×10^{-7} , 3.9984×10^{-8} , 1.3158×10^{-7} , 5.5487×10^{-10} , -4.968×10^{-8} , -1.1166×10^{-8} , and 6.1675×10^{-9} , respectively [21]. For a detailed analyze about the battery degradation cost please refer to [23]. The first term of (5) is the monetary cost imposed on customer k due to buying power x_k^h for his PEV and buying power l_k^h for his inelastic (non-PEV) appliances in price \mathcal{P}^h at slot h . The terms ϕ_k^s, ϕ_k^f in (5) define the battery degradation cost due to the variations of the stored energy (storage cost) and frequent fluctuations of charging/discharging power (operation cost) with model parameter $\gamma_k^{s1}, \gamma_k^{s2}, \gamma_k^{s3}$, and γ_k^f determined by the battery manufacturer (higher weights $\gamma_k^{s1}, \gamma_k^{s2}$, and γ_k^f encourage smaller variation with the constant per slot storage price γ_k^{s3}), respectively. Clearly, if the PEV power consumption does not change in some consecutive time slots, there is no operation cost ϕ_k^f for that slots [4].

III. PROBLEM FORMULATION

A. Tackling The Uncertainties

In general, the actual plug-in time α_k , plug-out time β_k , and the desired energy level (SOC) for the next trip E_k^d (i.e., the driving pattern) of each PEV k may not be available to its corresponding customer in advance. To tackle these uncertainty resources, the customer can collect the plug-in-plug-out and the desired energy demand historical data record of each PEV k . According to the collected driving pattern data, one can estimate the energy demand ($\hat{E}_k^d B_k^{cap}$) and the probabilities $\pi_{\alpha,k}(h)$ and $\pi_{\beta,k}(h)$ in which the PEV k becomes available and unavailable at each time slot h , respectively. The conditional probability $\pi_{\alpha,k}(h|\tau)$ that the PEV k becomes available in an upcoming time slot $h > \tau$, given that it has not become available until the current time slot τ , is [24]:

$$\pi_{\alpha,k}(h|\tau) = \frac{\pi_{\alpha,k}(h)}{1 - \sum_{h=1}^{\tau} \pi_{\alpha,k}(h)} \quad (7)$$

The conditional probability $\pi_{\beta,k}(h|\tau)$ that the PEV k becomes unavailable at slot h , given that it is still available at the current slot τ , can be drawn in a similar way. Bayes' theorem allows us to update prior beliefs on the probabilities $\pi_{\alpha,k}(h|\tau)$ and $\pi_{\beta,k}(h|\tau)$ at each slot h [25]. Due to the adaptive nature of our framework, from the point of view of Bayesian inference, one can estimate the values of α_k and β_k using the maximum a posteriori probability (MAP) estimate techniques [26]. In estimating the desired SOC E_k^d , if we underestimate the energy demand, high discomfort level

would impose on the PEV owner k and it is possible that the PEV's task remain unfinished. The smart conditional value-at-risk (CVaR)-based regularizer turns out to be capable of utilizing the PEVs capability intelligently for peak shaving and ancillary service effectively with limited underestimate risk [27]. To prevent the under-estimation (i.e., the energy deficit for the PEV), the customer k is charged with shortage price \mathcal{P}_{sh} (cents/kW), when its actual energy demand is higher than the estimated amount. Let $\Delta_k(E_k^d, \hat{E}_k^d) = \mathcal{P}_{sh}[E_k^d - \hat{E}_k^d]^+$ denote the penalty for energy deficit imposed on the customer k for one scheduling window, where \hat{E}_k^d is the estimated desired SOC. Under a given confidence level $\nu_k \in (0, 1)$ and the estimated energy demand \hat{E}_k^d , the value-at-risk (VaR) for PEV k is defined as the minimum threshold cost σ_k , for which the probability of energy shortage of PEV k being less than σ_k is at least ν_k . Due to the non-convexity, it is difficult to minimize the VaR. The CVaR is an alternative risk measure, which is convex and can be optimized using sampling techniques [27]. The CVaR, $CV_k^\nu(\hat{E}_k^d)$, for PEV k is defined as the expected value of the energy deficit cost, $\Delta_k(E_k^d, \hat{E}_k^d)$, when only the costs that are greater than or equal to the VaR, $V_k^\nu(\hat{E}_k^d)$, are considered [27]. That is,

$$\begin{aligned} V_k^\nu(\hat{E}_k^d) &= \min\{\sigma_k | \Pr\{\Delta_k(\cdot) \leq \sigma_k\} \geq \nu_k\}, \\ CV_k^\nu(\hat{E}_k^d) &= \mathbb{E}[\Delta_k | \Delta_k(\cdot) \geq V_k^\nu(\hat{E}_k^d)] \end{aligned} \quad (8)$$

It is possible to estimate the CVaR by adopting sample average approximation (SAA) technique [28]. Samples of the random variable E_k^d for PEV k can be observed from the historical record. Consider the set $\mathcal{E}_k = \{1, 2, \dots, E_k\}$ of E_k samples of the random variable E_k^d . Let $E_{k,n}^d$ denote the n^{th} sample of E_k^d for PEV k . The CVaR in (9) can be approximated by [29]:

$$\begin{aligned} CV_k^\nu(\hat{E}_k^d) &\approx \min_{\sigma_k} \tilde{C}_k^\nu(\hat{E}_k^d, \sigma_k), \\ \tilde{C}_k^\nu(\hat{E}_k^d, \sigma_k) &= \sigma_k + \sum_{n \in \mathcal{E}_k} \frac{[\Delta_k(E_{k,n}^d, \hat{E}_k^d) - \sigma_k]^+}{E_k(1 - \nu_k)} \end{aligned} \quad (9)$$

Under the given estimate \hat{E}_k^d , we can use the historical samples of the PEV's desired energy demand in each scheduling window to compute $\tilde{C}_k^\nu(\hat{E}_k^d, \sigma_k)$. Moreover, according to the MPC theory, we can use the information updated at each slot to improve the performance of the estimation [30].

B. The global optimization problem

From the aggregator point of view, the total load profile must be as flat as possible (valley filling) and the PEVs should be managed so that they provide maximum regulation capacity to reduce the power system losses and increase the efficiency. On the other hand, the customers only care about their payment. We provide a mechanism to make a compromise between the two. The global optimization problem to minimize

the customers' payments is given by:

$$\begin{aligned} \min_{\mathbf{x}_k \in \mathcal{X}_k} \sum_{k \in \mathcal{K}} \left(\sum_{h \in \mathcal{H}} \left((p^h + \lambda^h \sum_{k \in \mathcal{K}} x_k^h)(x_k^h + l_k^h) + \gamma_k^{s1}(x_k^h)^2 + \right. \right. \\ \left. \left. \gamma_k^{s2}|x_k^h| + \gamma_k^{s3} \right) + \sum_{h=2}^H \gamma_k^f (x_k^h - x_k^{h-1})^2 \right) \end{aligned} \quad (10)$$

This problem is a strictly convex due to its combination of strictly convex and linear functions [31]. We can solve problem (10) centrally by the aggregator using well known convex programming techniques. However, due to the mentioned drawbacks of the centralized solutions (in Section I-B), we seek to provide a decentralized solution. In Section V of [5] it is shown that under some conditions, the decentralized game theoretic solution (Nash equilibrium) coincides with the optimal global centralized solution. In the decentralized strategy, each customer is aware only of his personal objective function and constraints. With some manipulation the global problem (10) is rewritten as follows:

$$\begin{aligned} \min_{\mathbf{x}_k \in \mathcal{X}_k} \sum_{k \in \mathcal{K}} \sum_{h \in \mathcal{H}} \left(\lambda^h \sum_{k \in \mathcal{K}} x_k^h \right) (x_k^h + l_k^h) + \sum_{k \in \mathcal{K}} \sum_{h \in \mathcal{H}} p^h (x_k^h + l_k^h) + \\ \sum_{k \in \mathcal{K}} \sum_{h \in \mathcal{H}} \gamma_k^{s1}(x_k^h)^2 + \gamma_k^{s2}|x_k^h| + \gamma_k^{s3} + \sum_{h=2}^H \gamma_k^f (x_k^h - x_k^{h-1})^2 \end{aligned} \quad (11)$$

The challenging term for decentralizing the problem is the first term which shows that the optimal behavior of each customer depends on how he/she and all other customers schedule their consumptions. So, due to the dynamic price component $\lambda^h \sum_{k \in \mathcal{K}} x_k^h$, the customers' actions are coupled to each other through an energy cost sharing model. The most suitable solution concept to tackle such energy cost sharing model is achieved through *game theory*. Each game $\Gamma(\mathcal{K}, \mathbf{x}_k, P_k(\mathbf{x}_k, \mathbf{x}_{-k}))$ is defined by three components as: *Players*: All the registered customers in set \mathcal{K} . *Strategies*: Charging/discharging schedules \mathbf{x}_k of each player $k \in \mathcal{K}$. *Payoffs*: Local objective function $P_k(\mathbf{x}_k, \mathbf{x}_{-k})$ that each player k seeks to maximize it. Let's define the local cost function associated with each single customer $k \in \mathcal{K}$ as follows:

$$\begin{aligned} \min_{\mathbf{x}_k \in \mathcal{X}_k} J_k(\mathbf{x}_k, \mathbf{x}_{-k}) &= \sum_{h \in \mathcal{H}} \left(\left(\lambda^h (x_k^h + \sum_{j \in \mathcal{K}/k} x_j^h) \right) (x_k^h + l_k^h) + \right. \\ &\left. p^h (x_k^h + l_k^h) + \gamma_k^{s1}(x_k^h)^2 + \gamma_k^{s2}|x_k^h| + \gamma_k^{s3} \right) + \sum_{h=2}^H \gamma_k^f (x_k^h - x_k^{h-1})^2 \end{aligned} \quad (12)$$

where \mathbf{x}_{-k} denotes all the customers' action other than that of customer k . The payoff function of customer k becomes $P_k(\mathbf{x}_k, \mathbf{x}_{-k}) = -J_k(\mathbf{x}_k, \mathbf{x}_{-k})$. The dynamic price component is decomposed accordingly to customer k 's dynamic cost and the other customers ($j \in \mathcal{K}/k$) as $\lambda^h (x_k^h + \sum_{j \in \mathcal{K}/k} x_j^h)$.

Theorem 1. *An optimal solution of (11) provides a valley filling load profile, minimizes the load variance, minimizes the power system losses, and maximizes the load factor.*

Proof. See Appendix A. □

In view of Theorem 1, as the proposed framework has the lowest load variance and results in a valley filling charging profile, there is no need to be concerned about additional constraint that may refrain from equipment's overloading.

Theorem 2. *An optimal on-off strategy with maximum charge rate x_k^r for convex minimization problem (12) maximizes the regulation capacity.*

Proof. See Appendix B. \square

Based on Theorem 2, the charging control should be on or off at the maximum charging rate to maximize regulation capacity. Thus, there is no need to be concerned about the optimal charging rate and constraints such as $x_k^{min} \leq x_k^h \leq x_k^{max}$. The only important thing is how to schedule the charging sequence. So, the power consumption is constrained to be off ($x_k^h = 0$) or on ($x_k^h = x_k^r$) and the feasible set \mathcal{X}_k for each PEV owner $k \in \mathcal{K}$ is modified as follows:

$$\tilde{\mathcal{X}}_k = \{\mathbf{x}_k \mid (1), (2), \text{ and } x_k^h \in \{0, x_k^r\}\} \quad (13)$$

where x_k^r is the rated power at which the PEV k is charged (> 0) or discharged (< 0).

IV. NASH FOLK REGRET-BASED STRATEGY

The system parameters, such as, the price signal, the base loads, the deriving patterns, and the customer's decisions can have unpredictable changes. So, there is always a difference between the real world and the scheduled energy consumptions. This difference results in a sub-optimal schedule, even with deploying the MPC and stochastic methods. Assuming perfect knowledge of the parameters' behavior throughout the scheduling horizon is very strong. Therefore, we adopt the optimistic mirror descent (OMD) regret-based algorithm to take into account the long term behavior of these unpredictable changes. The iterative minimization of problem (12) for each $k \in \mathcal{K}$ based on the OMD algorithm becomes:

$$\begin{aligned} \psi_k^{t+1} &= \nabla L_k^{-1}(\nabla L_k(\psi_k^t) - \eta_k \nabla J_k^t(\mathbf{x}_k^t, \mathbf{x}_{-k}^t)) \\ \mathbf{x}_k^{t+1} &= \arg \min_{\mathbf{x}_k \in \tilde{\mathcal{X}}_k} \eta_k \mathbf{x}_k^\top M_k^{t+1} + D_{L_k}(\mathbf{x}_k, \psi_k^{t+1}) \end{aligned} \quad (14)$$

where superscript t denotes the problem at hand in day t , η_k is a penalty term, L_k is any 1-strongly convex function, and M_k^{t+1} denotes prediction for the gradient of the cost function J_k^{t+1} . For example, we can select the prediction M_k^{t+1} as the average of the gradients of the cost functions for the previous days, i.e., $M_k^{t+1} = 1/t \sum_{v=1}^t \nabla J_k^v(\mathbf{x}_k^v, \mathbf{x}_{-k}^v)$ [7]. This is because, it is expected that the system parameters have high correlations with the previous days (the customers have daily habits, go/come back to/from the work at similar hours, etc). So, the OMD algorithm uses the information of the previous scheduling horizons (days) to improve the long-term performance. Let's consider the customer k regret after T days as:

$$Rg_k(T) := \sum_{t=1}^T J_k^t(\mathbf{x}_k^t, \mathbf{x}_{-k}^t) - \min_{\mathbf{x}_k \in \tilde{\mathcal{X}}_k} \sum_{t=1}^T J_k^t(\mathbf{x}_k, \mathbf{x}_{-k}) \quad (15)$$

In fact, applying such regret-based OMD algorithm provide a data-base of action-results to the customer k by which the customer can infer each of his actions results to what cost (how

much regret). This approach can be seen as a way of adding prior knowledge about the sequence within the paradigm of online learning [32].

Proposition 1. *Iteration (14) converges in the sense that, for any $\mathbf{x}_k^* \in \tilde{\mathcal{X}}_k$ we have:*

$$Rg_k(T) \leq \frac{1}{\eta_k} Y_k + \frac{\eta_k}{2} \sum_{t=1}^T \|\nabla J_k^t(\mathbf{x}_k^t, \mathbf{x}_{-k}^t) - M_k^t\|_*^2 \quad (16)$$

where

$$Y_k := \max_{\mathbf{x}_k \in \tilde{\mathcal{X}}_k} L_k(\mathbf{x}_k) - \min_{\mathbf{x}_k \in \tilde{\mathcal{X}}_k} L_k(\mathbf{x}_k)$$

Proof. See Proof of Lemma 2 in [33]. In particular, if η_k is chosen as $O(1/\sqrt{T})$, then the average regret, namely, $Rg_k(T)/T$ converges to zero as $T \rightarrow \infty$ [7]. \square

Now, we need to provide a coordinated charging mechanism between the PEVs to result in a public social optimal solution to the global problem (11). The main challenge to provide a fully decentralized solution to (11) is that each customer k requires to know the aggregate PEVs power consumption $\sum_{k \in \mathcal{K}} x_k^h = x_k^h + \sum_{j \in \mathcal{K}/k} x_j^h$ for each slot $h \in \mathcal{H}$ to solve (12). A common state of the art solution to this, is that each PEV submit its action profile to all other PEVs (either directly, or through the aggregator). According to this mechanism, each PEV k independently executes the Nash equilibrium regret-based Algorithm 1 to achieve optimal coordinated charging (See [34] for optimality and convergence proof). This method

Algorithm 1 Decentralized Nash Equilibrium Strategy (DNES): executed by each PEV k

- 1: **I. Initialization:** Take arbitrary action x_k^0 , M_k^0 , and ∇J_k^0 .
 - 2: **II. Repeat** for $t = 1, 2, \dots$
 - 3: Set $h = 0$, and take parameters of previous day $t - 1$ as initial values.
 - 4: **Repeat according to some order**
 - 5: Receive the last update of the other PEVs actions (\mathbf{x}_j^t in $j \in \mathcal{K}/k$).
 - 6: Run OMD algorithm (14) and determine \mathbf{x}_k^{t+1} .
 - 7: **If** \mathbf{x}_k^{t+1} changes compared to current schedule, **Then**, set \mathbf{x}_k^{t+1} as the new solution and broadcast it.
 - 8: **Until convergence**
-

requires very low-latency communication, is not robust to noise and node/link failures, and jeopardizes the customers' privacy security. More importantly, for a large power system, it is probable that some customers, deliberately or inadvertently, do not communicate with other or do not act rationally. We can model the noisy data such as the customers are not so rational and reliable. In this situation the classical games fail to reach the Nash equilibrium. So, we have used a behavioral repeated game in this paper. Our game can be modeled as an infinite game due to the fact that the energy scheduling must be repeated at each hour or at each day. To tackle this issue, we present a fully decentralized robust collaborative method based on the *Nash Folk* theorem [35]. The original Folk theorem concerned the payoffs of all the Nash equilibria of an infinitely repeated game. This theorem helps to identity the deviator players (PEVs), and punish them with locally and private message exchange between the neighbors. There are two kinds of deviations (disruptions) in action or in communication. If a player either adopt any strategy other than the provided optimal equilibrium strategy of Algorithm

1 or sends a spurious information to his neighbors, he is a deviator. Let's define the PEV k 's neighborhood \mathcal{N}_k as the set of all PEVs with ability to communicate with PEV k . The key to characterize a scalable game framework lies in determining under which conditions can information be diffused through the grid by the neighbors. In this way, we propose Algorithm 2 considering the following proposition.

Proposition 2. *The necessary and sufficient conditions for a full Nash Folk theorem to hold and to enable deviator identification by the deviator's neighbors are as follows:*

- 1) Condition P: For each player $k \in \mathcal{K}$, each neighbor $j \in \mathcal{N}_k$, any actions $\mathbf{x}_j, \hat{\mathbf{x}}_j \in \tilde{\mathcal{X}}_j$, $\mathbf{x}_k \in \tilde{\mathcal{X}}_k$ and $\mathbf{x}_{\mathcal{N}_k/j} \in \tilde{\mathcal{X}}_{\mathcal{N}_k/j}$, we must have $J_k(\mathbf{x}_k, \mathbf{x}_j, \mathbf{x}_{\mathcal{N}_k/j}) \neq J_k(\mathbf{x}_k, \hat{\mathbf{x}}_j, \mathbf{x}_{\mathcal{N}_k/j})$.
- 2) Condition N: For each player k , for any neighbors $j, i \in \mathcal{N}_k$ such that $j \neq i$, there exists $\ell \in \mathcal{N}_j/i \Delta \mathcal{N}_i/j$, such that there is a path from ℓ to k which passes through neither j nor i .

where \mathcal{N}_k/j is the set of player k 's neighbors other than neighbor j with action set $\tilde{\mathcal{X}}_{\mathcal{N}_k/j}$ and action profile $\mathbf{x}_{\mathcal{N}_k/j}$. $\ell \in \mathcal{N}_j/i \Delta \mathcal{N}_i/j = (\mathcal{N}_j/\{i, \mathcal{N}_i\}) \cup (\mathcal{N}_i/\{j, \mathcal{N}_j\})$ and implies that player ℓ is a neighbor of j or i , but not of both. Condition P implies that the action of a players always affects his neighbors' payoffs, so, its behavior is learnable (deviation is detectable). A 3-connected² network in which players have different neighbors, i.e., for any pair $(j, i) \in \mathbb{N}^2, \mathcal{N}_j/i \neq \mathcal{N}_i/j$, satisfies Condition N (see [36] for detailed information and proof about Proposition 2). In the proposed framework in Algorithm 2, the payoff of a customer depends only on his own action and the actions of his neighbors. This algorithm is robust to (deliberately or inadvertently) disruption in the sense that it has ability to identify the guilty customers with the local private information exchange as follows:

Communication protocol CP: Let's denote by $\theta = \inf\{t \geq 1 : \exists k \in \mathcal{K} \text{ s.t. } (\mathbf{x}_k^t \neq \mathbf{x}_k^{*t} \text{ or } \exists j \in \mathcal{N}_k \text{ s.t. } m_k^t(j) \neq \emptyset)\}$, the first stage at which a player starts deviating, where \mathbf{x}_k^{*t} is the customer k 's equilibrium strategy at stage t and $m_k^t(j) = \{k, \dots, i\}$ means that a deviation report has been sent from player k to player j at stage t declaring players $\{k, \dots, i\}$ as the innocent players. When each customer $k \in \mathcal{K}$ detects a deviation ($\theta \neq \emptyset$), that is, either he observes a change in his payoff or he receives $m_j^t(k) \neq \emptyset$ from his neighbor, then he sends the sets of innocent players to his neighbors. This set is computed as follows:

- If he has observed a change in his payoff (the deviation is in action), then he clears all the customers that are not his neighbors. That means, the deviator must be one of his neighbors. If $\theta \neq \emptyset$ and he does not observe a change in his payoff (he has received $m_j^t(k)$ from some $j \in \mathcal{N}_k$), he clears all his neighbors.
- The set of innocent customers then is updated using the sets received by his neighbors, i.e., if neighbor $j \in \mathcal{N}_k$ sends his set of innocents $\{j, i, m\}$, customer k adds

²Graph G with vertex V is called n -connected if $|V| \geq n$ and $G - X$ is connected for every set $X \subseteq V$ with $|X| < n$, where $G - X$ represents the graph where all nodes in X have been removed (and the corresponding links). Simply, a graph is n -connected if any two of its nodes can be joined by n independent paths.

customer i and m to his own set of innocents. Note that player k cannot clear player j . If j is the deviator, then i and m are cleared automatically. Otherwise, j is performing the protocol obediently, so his information is true and i and m are really innocent. Therefore, by this mechanism the information “ j claims that i and m are innocent” is not manipulable by player j .

- At the end of the protocol, if at least one customer $k \in \mathcal{N}_\ell$ says that customer $\ell \in \mathcal{N}_k$ is not cleared, then customer ℓ is known as the deviator. Otherwise, there is no action deviation and deviation is in communication.

The adopted communication protocol is such that, if a customer deviates in communication but not in action, then all customers keep playing the equilibrium actions and payoffs are not affected. Moreover, when there is an action deviation, only the deviator's neighbors have to identify him in order to punish him.

Punishment: The (independent) minmax level of player ℓ to be punished by his neighbors is as follows:

$$v_\ell = \min_{\mathbf{x}_{\mathcal{N}_\ell} \in \prod_{j \in \mathcal{N}_\ell} \tilde{\mathcal{X}}_j} \max_{\mathbf{x}_\ell \in \tilde{\mathcal{X}}_\ell} P_\ell(\mathbf{x}_\ell, \mathbf{x}_{\mathcal{N}_\ell}) \quad (17)$$

where $P_\ell(\mathbf{x}_\ell, \mathbf{x}_{\mathcal{N}_\ell}) = -J_\ell(\mathbf{x}_\ell, \mathbf{x}_{\mathcal{N}_\ell})$ and $J_\ell(\mathbf{x}_\ell, \mathbf{x}_{\mathcal{N}_\ell})$ is obtained from (12) by replacing $\sum_{j \in \mathcal{N}_\ell} x_j^h$ for $\sum_{j \in \mathcal{K}/\ell} x_j^h$.

Algorithm 2 Fully Decentralized Nash Folk Strategy (FD-NFS): Executed by each PEV k

- 1: **I. Initialization:** Take arbitrary action x_k^0 , M_k^0 , and ∇J_k^0 .
- 2: **II. Repeat for** $t = 1, 2, \dots$
- 3: Set $h = 0$, and take parameters of previous day $t - 1$ as initial values.
- 4: **Repeat according to some order**
- 5: **Action Phase:** Receive the last update \mathbf{x}_j^t of each PEV in neighborhood $j \in \mathcal{N}_k$.
- 6: Run OMD algorithm (14) and determine \mathbf{x}_k^{t+1} .
- 7: **Communication Phase:** If \mathbf{x}_k^{t+1} changes compared to current schedule, **Then** set \mathbf{x}_k^{t+1} as the new solution and broadcast it to the neighbors.
- 8: **If** no deviation is detected, **Then** sends a blank message $m_k^{t+1}(j) = \emptyset$ to his neighbors $j \in \mathcal{N}_k$.
- 9: **Else** run the communication protocol CP until the deviator ℓ is identified.
- 10: **Punishment Phase:** If the deviator ℓ is neighbor with k , **Then** player k plays minmax strategy (17) against his neighbor ℓ .
- 11: **Until convergence**

To provide an insight about how the customers interact with each other and apply the Nash Folk strategy, the block diagram of the proposed framework is presented in Fig. 1.

Theorem 3. *The developed Nash Folk repeated game in Algorithm 2 converges to an optimal unique solution of global minimization problem (11) at Nash equilibrium. In the Nash equilibrium solution, no player can increase his payoff by switching unilaterally to an alternative strategy.*

Proof. See Appendix C. □

Remark 1. *For improving the short term performance of the proposed mechanism, we can use an event-triggered MPC strategy with optimal tuning t and run the repeated game once needed (e.g., a change is occurred) [30].*

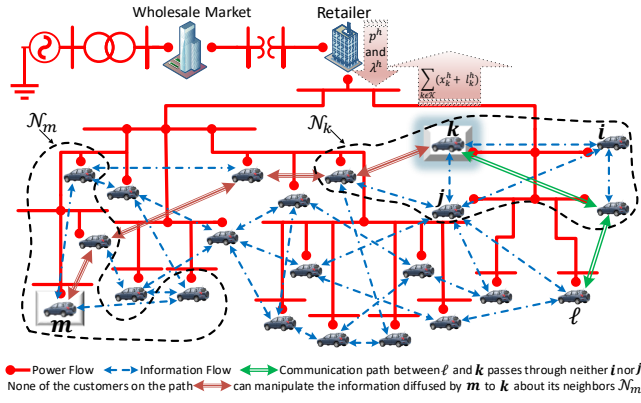


Fig. 1. The block diagram of applying the Nash Folk strategy on the considered smart micro-grid model.

V. NUMERICAL RESULTS

A. Simulation Scenarios and System Parameters

In the adopted setup for numerical simulations, we have considered a smart micro-grid including one aggregator procuring power for $K = 10$ customers within scheduling horizon $H = 24$ for each $h = 1$ hour. The baseline price signal is adopted from the real case day ahead hourly locational marginal price (LMP) values determined by Pennsylvania-New Jersey-Maryland Interconnection (PJM) electricity market for 2017/08/01 [37]. Each customer has one PEV with ideal charge/discharge efficiency and their base loads is adopted randomly with some correlation with the baseline price signal. The customers' electric vehicles are assumed to randomly have different specifications (e.g., $12 \leq B_k^{cap} = E_k^{max} \leq 14$ kWh, $E_k^{min} = 0$, $x_k^r = 4$ kW/h) and preferences (e.g., \mathcal{H}_k). We also assumed that the initial energy level of all the PEVs is zero and $12 \leq E_k^d \leq 14$ kWh. Without loss of generality, for the considered system topology (graph), it is assumed that each customer has three neighbors (e.g., $\mathcal{N}_1 = \{3, 4, 5\}$, $\mathcal{N}_2 = \{4, 5, 6\}$, $\mathcal{N}_3 = \{1, 5, 6\}$, $\mathcal{N}_4 = \{1, 2, 6\}$, $\mathcal{N}_5 = \{1, 2, 3\}$, $\mathcal{N}_6 = \{2, 3, 4\}$, $\mathcal{N}_7 = \{6, 8, 9\}$, $\mathcal{N}_8 = \{5, 7, 10\}$, $\mathcal{N}_9 = \{2, 7, 10\}$, $\mathcal{N}_{10} = \{3, 8, 9\}$). In the proposed results, **PP** and **PT** denote dynamic price-participant and static price-taker scenarios, respectively. The grid topology with 3-connected graph is denoted by 3 and 2-connected graph is denoted by number 2, and symbol **B** denotes a communication protocol in which each customer needs to send his base load to the neighbors. Further, **on/off (continuous)** denotes the scenario in which the PEVs are allowed only to charge/discharge with fixed (modulated) charging rate, and **V2G** denotes the scenario in which the PEVs can sell electricity back to the grid.

B. Performance Comparison

For each of the algorithms (DNES and FDNFS) formulated in the previous section, the convergence behavior is compared for different price signals, communication protocols, and grid topologies in Fig. 2. As seen in this figure, deploying both of the algorithms with different assumptions the system converges very fast (only one iteration per customer is required) and result in much lower total grid payment compared with when there is not any scheduling program. In all the results

the proposed FDNFS is as efficient as the DNES method (which the dominant method in the literature), while it is obvious that our novel FDNFS method is more robust, more reliable, require low communications, and scalable. Also this results indicate that there is no need for customers to share their base non-PEV loads (as long as the price signal is well designed) and they are free to adopt 2, 3, or more neighbors (as long as the grid topology is a connected graph). As in the price-participant scenarios, the consumption patterns affect the price signal, there is some oscillation in the convergence process around the optimal point. This is because, as much as the customers shift their PEVs consumption to time slots with lower prices, the spot price increases at those slots to prevent creation of sub-peaks [18]. The performance of the centralized method in the price taking (PT) scenario shows that the aggregate system payment is the lowest. We must note that when there is too many customers in the system, the centralized method is not practical as it imposes a huge computational burden to the system, the privacy of the customers are put into danger, and the algorithm is not reliable and robust to failures. However, in another simulation scenario with dynamic pricing (Centralized-PP), we can see that performance of the centralized method is worse than the decentralized methods. This is because in the decentralized method, at each iteration the customers update the price function and effectively can predict the effect of their actions on the price behavior.

However, according to Fig. 2, when the customers are assumed to be price-taker, they incur more cost compared with those by price-participant customers. This is because when the customers are price-participant, shifting the PEVs' load to lower price slots helps flatten and lower the price signal. This behavior is illustrated in Fig. 3. As shown in Fig. 3(a), without the charge scheduling, the PEVs' load demand deteriorates the system performance as results in higher peak demand. Accordingly, when our method (the FDNFS algorithm) is deployed, the peak demand is reduced and the PEVs' charging profiles is shifted to the time slots with lower price. This behavior results in a valley filling pattern. Fig. 3(b) illustrates the deference between dynamic-PP and static-PT real time pricing. In the case of static pricing, the customers' behavior does not affect the price signal. So, the spot price is coordinated with the baseline price to reduce the peak demand as much as possible. On the other hand, when dynamic real time pricing is performed, shifting the PEVs' demand to low price time slots results in transferring the spot price to those slots. This behavior has two benefits; preventing creation sub-peaks in slots with low prices and reducing the customers' payments (see Fig. 4).

To make incentive for all the PEV owners to participate in a scheduling program they must be sure to acquire some monetary benefit of it. From Fig. 4, we can see that all the PEV owners can reduce their payments by participating in our mechanism. As previously mentioned, for all the simulations the customer's payment with dynamic-PP pricing is lower than other scenarios (no schedule and static-PT pricing). Moreover, it is obvious that by V2G protocol, the customers can charge their PEVs in low price slots and sell it at high price slots (see Fig. 5), resulting in minimum possible payment.

As an example, customer 4's load pattern in four scenarios are given in Fig. 5. Two charging models (**on/off** and **continuous**) without ability of selling electricity back to the grid are shown in Fig. 5(a). In both models the customer has charged his PEV in slots with low base load demand (which coincides with low price slots). As is clear from Fig. 5(a), in the on/off charging mode, the PEV is working on only three time slots, while **continuous** charging occupies six slots. As we proved in Appendix B, this result shows that the on/off charging strategy provides more regulation capacity. The charging profile for customer 4 under V2G protocol is shown in Fig. 5(b). We can see that the customer has charged his PEV with more power than he needs in slots with low prices and sell back the excess energy in slots with high prices to reduce his payment. However, due to the frequent charge/discharge battery cycles, this procedure increases the battery degradation cost (see Fig. 6).

There is always a trade-off between reducing the payment by frequently charging and discharging the storage devices and reducing the battery degradation cost by keeping constant the battery SOC. Comparison between Fig. 4 and Fig. 6 demonstrates this trade-off. The results of Fig. 4 show that the customers have minimum possible payment with the V2G protocol. On the other perspective, Figs. 6(a)-(b) illustrate that the V2G protocol results in maximum possible battery degradation cost.

The long-term performance of the proposed framework is analyzed in Figs. 7(a)-(b). The aggregate behavior of 10 customer in the presence of uncertainty about parameters α_k , β_k , E_k^d , and p^h are demonstrated in Fig. 7(a). As we can see, at time slots with high price (which is mostly incurred by high base load) all the customers try to discharge the battery of their PEVs to make profit. They also charge their PEVs at slots with low price to pay less. However, the most(least) PEV power charging(discharging) does not exactly coincides with the lowest(highest) price due to inaccuracy in stochastic parameters estimation. Moreover, the convergence behavior of iteration (14) for customer 4 is depicted in Fig. 7(b) by regret criterion (15). The result denotes that at the first days the regret level is high, while as the learning capability of the customer increases, the regret level goes down and converges.

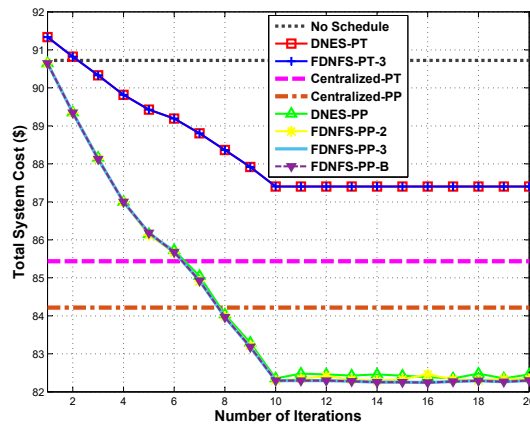


Fig. 2. Algorithms 1-2 convergence behavior comparison in different situations.

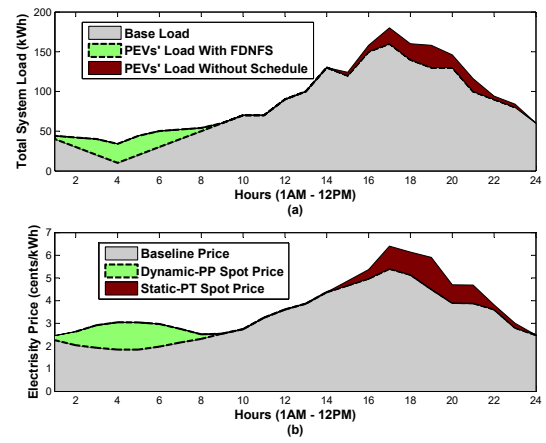


Fig. 3. Total system load and price patterns: (a) PEV's energy consumption behavior and (b) Difference between dynamic and static pricing.

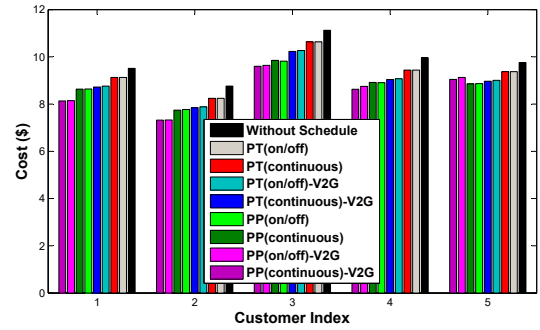


Fig. 4. Total cost imposed on each customer in different scenarios.

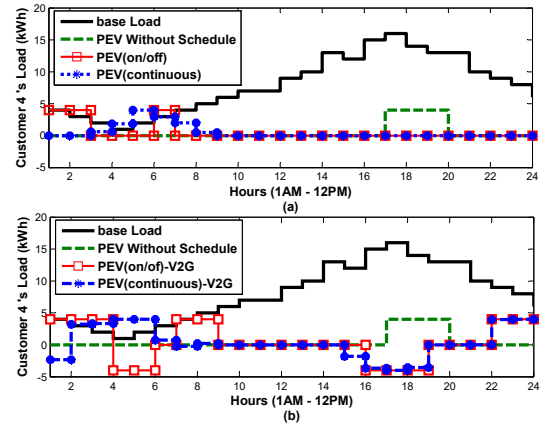


Fig. 5. Customer 4's load profile: (a) Without ability to selling electricity back and (b) With ability to selling electricity back (V2G).

VI. CONCLUSIONS

Efficient charging scheduling of PEVs requires robust, decentralized and incentive-based coordination mechanism. None of the proposed PEV charging/discharging schedules in the literature is robust fully decentralized. To address these concerns, we proposed a computationally efficient “fully” decentralized game-based charging mechanism coordinated by a real time dynamic pricing method using the Nash Folk strategy. Due to the private and local communication between the neighbors this framework is robust to link/node failures and

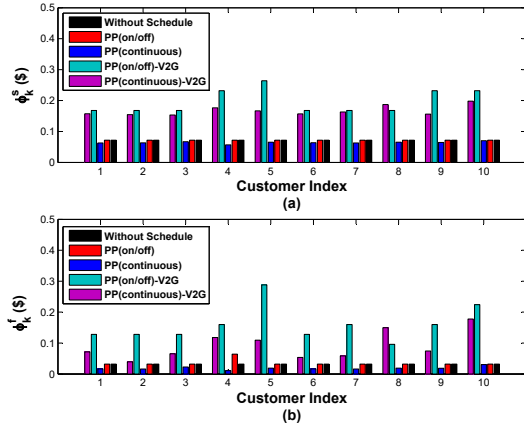


Fig. 6. Battery cost imposed to each customer: (a) Battery degradation cost due to the amount of charging/discharging power and (b) Battery degradation cost due to the fluctuations of charging/discharging power.

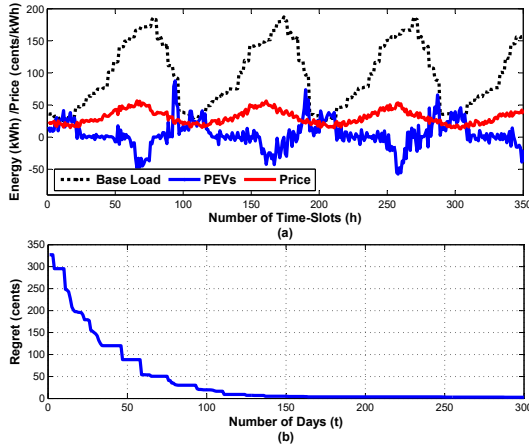


Fig. 7. Long-term performance of the proposed framework under uncertainty of the deriving pattern and the baseline-price: (a) aggregate consumption pattern (b) regret convergence of customer 4.

noisy information too. Our pricing and charging schedule takes into account the PEV's battery specification and degradation cost, is completely general, and can work with any scenario with some manipulation. The MPC and regret matching techniques were used with Bayesian inference and the CVaR to cope with uncertainties and improve the short/long-term performance of the mechanism. Further, for the first time, we illustrated that any cheating by the players or any cyber-attack can be identified by the proposed framework. We proved that our method minimizes the load variance and power system losses, and provides maximum regulation capacity and the load factor. Finally, we proved and demonstrated by numerical results that our mechanism works efficiently in different scenarios and results in total valley filling profile and minimum possible payment for the customers.

APPENDIX

A. Proof of Theorem 1

The proof of Theorem 1 consists of four parts. In the first part we need to show that an optimal solution of (11) provides a valley filling load profile. To this end we need following proposition

Proposition 3. *Minimizing an strictly convex global function $\sum_{k \in \mathcal{K}} U(\mathbf{x}_k) : \mathbb{R} \rightarrow \mathbb{R}$ results in a valley filling charging profile $\sum_{k \in \mathcal{K}} x_k^h = [A - \sum_{k \in \mathcal{K}} l_k^h]^+, \forall h \in \mathcal{H}$, where $[a]^+ = \max\{0, a\}$ with some $A \in \mathbb{R}$.*

Proof. In [6], it is proved that a valley-filling charging profile is optimal (Property 1) and an optimal solution exist, if $\mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_K \neq \emptyset$, i.e., \mathcal{X} is nonempty (Property 2). The feasible set \mathcal{X}_k of each PEV k 's charging profile is obviously nonempty, compact, and convex. So, the global feasible set is also nonempty, compact, and convex [31].

Let $\mathbf{X} \triangleq [\mathbf{x}_1, \dots, \mathbf{x}_K]^T \in \mathbb{R}^{K \times H}$, $\mathbf{x} = \mathbf{x}_1 + \dots + \mathbf{x}_K \in \mathbb{R}^{1 \times H}$ denote the charging profile and the aggregate charging profile of all the PEVs.

Definition 1. *Two feasible charging profiles \mathbf{X} and $\hat{\mathbf{X}}$ are equivalent, denoted as $\mathbf{X} \cong \hat{\mathbf{X}}$, if their aggregate charging profiles are the same, i.e., $\mathbf{x} = \hat{\mathbf{x}}$.*

It is inference from Theorem 1 in [6] that for optimization problem (12), optimal charging profile is in general not unique. So, let denote the set of optimal charging profiles with an arbitrary optimal charging profile \mathbf{X}^o as $\mathcal{O} := \{\mathbf{X} \in \mathcal{X} \mid \mathbf{X} \cong \mathbf{X}^o\}$. Moreover, it is shown in [6] that the set \mathcal{O} of optimal charging profiles does not depend on the choice of $U(\mathbf{X})$. That means, if \mathbf{X}^o is optimal with respect to a strictly convex $U^1(\mathbf{X})$, then \mathbf{X}^o is optimal with respect to any strictly convex $U^n(\mathbf{X})$.

Now, we consider an arbitrary valley-filling charging profile \mathbf{X}^v . From property 1 of [6], this charging profile is optimal to any strictly convex function, specially to (11). As the objective function of problem 12 and its global objective function (11) are also strictly convex, each global optimal solution \mathbf{X}^o to (11), is an equivalence class with valley-filling charging profile \mathbf{X}^v . From Definition 1 the aggregate charging profiles of \mathbf{X}^v and \mathbf{X}^o are the same. So, we can conclude that the optimal charging profile \mathbf{X}^o generates a valley-filling profile and the proof of Proposition 3 (the first part of the proof of Theorem 1) is complete. \square

Next, we proceed to show that an optimal solution of (11) minimizes the load variance (the second part of the proof of Theorem 1). Note that as the set \mathcal{O} of optimal charging profiles does not depend on the choice of $U(\mathbf{X})$, without loss of generality we assume that our objective function is $U(\mathbf{X}) = \sum_{k \in \mathcal{K}} \sum_{h \in \mathcal{H}} (x_k^h)^2$. The load variance seen at the retailer side is defined as follows:

$$\sigma^2 \triangleq \frac{1}{H} \sum_{h \in \mathcal{H}} (X^h - L_{avg})^2, \quad L_{avg} = \frac{1}{H} \sum_{k \in \mathcal{K}} \sum_{h \in \mathcal{H}} (x_k^h - l_k^h) \quad (18)$$

where $X^h = \sum_{k \in \mathcal{K}} x_k^h$ is the aggregate PEVs' load at slot h . Function (18) can be expanded as follows:

$$\sigma^2 \triangleq \frac{1}{H} \sum_{h \in \mathcal{H}} (X^h)^2 - 2L_{avg} \frac{1}{H} \sum_{h \in \mathcal{H}} X^h + \frac{1}{H} \sum_{h \in \mathcal{H}} L_{avg}^2 \quad (19)$$

where the second term must be constant (equal to $\sum_{k \in \mathcal{K}} E_k^d / (\eta_k B_k^{cap})$) due to constraint (2). Since L_{avg} is constant, the last term is constant likewise. Therefore, minimizing the load variance is equivalent to minimizing the first term of (19), i.e., our objective function $\sum_{h \in \mathcal{H}} (X^h)^2$. So, the optimal charge profiles which minimize our global problem (11), result in the minimum load variance.

In the third part of the proof of Theorem 1, we need to show that an optimal solution of (11) minimizes the power system losses. According to the Theorem 2 in [3], for a fixed energy demand in scheduling horizon H , minimizing the feeder energy losses is equivalent to minimizing load variance if the feeder is a single branch with small voltage fluctuations, which completes the third part of the proof of Theorem 1.

Finally, we proceed to fourth part of the proof of Theorem 1. As the PEV consume much power, in Section II-B, we assumed that the PEVs penetration into the future power system is large enough to affect the spot electricity price. So, it is reasonable to assume that $\sum_{k \in \mathcal{K}} E_k^d / (\eta_k B_k^{cap})$ is larger enough compared with $\sum_{k \in \mathcal{K}} \sum_{h \in \mathcal{H}} l_k^h$. Let $I_{m,b}$, V denote maximum current of the aggregate base loads and the system voltage, respectively. According to the Theorem 3 in [3], if $\sum_{k \in \mathcal{K}} \sum_{h \in \mathcal{H}} (x_k^h + l_k^h) \geq I_{m,b} VT$ hold, minimizing the load variance is equivalent to maximizing the load factor and the proof of Theorem 1 is complete.

B. Proof of Theorem 2

A vehicle is considered to provide regulation service and can make profit when being idle. So, the more the PEV is in idle mode, the more it can provide the regulation capacity. On the other hand, when a vehicle is charging, it has to pay for purchasing the power from the grid. Thus, the profit acquired by PEV owner $k \in \mathcal{K}$ is defined as follows (for simplicity and without loss of generality we do not consider the degradation cost and we drop subscript k from the following notations):

$$G(\mathcal{H}_c, r^h) \triangleq \sum_{h \in (\mathcal{H} - \mathcal{H}_c)} \mathcal{P}_R^h - x^{max} \sum_{h \in \mathcal{H}_c} \mathcal{P}^h r^h \quad (20)$$

where \mathcal{H}_c is the set of time slots in which PEV k is in charging mode, \mathcal{P}_R^h is the regulation price paid to the PEV for providing the regulation capacity at slot h , x^{max} is the maximum possible charging rate, and $0 \leq r^h \leq 1$ is the charging rate. Rewriting the above equation with respect to the control variables r^h and \mathcal{P}^h yields:

$$G(\mathcal{H}_c, r^h) \triangleq \sum_{h \in \mathcal{H}} \mathcal{P}_R^h - \underbrace{\sum_{h \in \mathcal{H}_c} (\mathcal{P}^h r^h x^{max} + \mathcal{P}_R^h)}_{H(r^h)} \quad (21)$$

where the first term is constant and \mathcal{P}_R^h , \mathcal{P}^h are given by the aggregator. So, to maximize the profit we need minimize the second term ($H(r^h)$) of (21). Let's assume that there exist optimal r_1^h and r_2^h such that $0 < r_1^h, r_2^h < 1$. It follows then that we have:

$$\begin{aligned} H(r_1^h) + H(r_2^h) &= \mathcal{P}_1^h r_1^h x^{max} + \mathcal{P}_{R,1}^h + \mathcal{P}_2^h r_2^h x^{max} + \mathcal{P}_{R,2}^h \\ &= r_1^h (\mathcal{P}_1^h - \mathcal{P}_2^h) + D \mathcal{P}_2^h + \mathcal{P}_{R,1}^h + \mathcal{P}_{R,2}^h \end{aligned} \quad (22)$$

As we seek to determine a charging rate which results in the maximum regulation capacity, we can assume $D = r_1^h + r_2^h$ to be constant. Therefore, (22) becomes a first-order equation with respect to r_1^h and the extrema occur only at both ends of the range of r_1^h . That is, (22) has a minimum at either of $r_1^h = 0$ or $r_1^h = 1$ from (21), which contradicts the assumption that $0 < r_1^h, r_2^h < 1$. Therefore, r^h is unique and should be either 0 or 1 to minimize $H(r^h)$. Since we assumed that \mathcal{P}_R^h ,

\mathcal{P}^h are known (given by the aggregator), minimizing $H(r^h)$ results in minimizing \mathcal{H}_c and consequently maximize the time slots in which the PEV is in idle mode.

C. Proof of Theorem 3

As our objective function $f_k(\mathbf{x}_k)$ is strictly convex, each payoff function $P_k(\mathbf{x}_k, \mathbf{x}_{-k})$ is strictly concave with respect to \mathbf{x}_k [31]. Therefore, game $\Gamma(\mathcal{K}, \mathbf{x}_k, P_k(\mathbf{x}_k, \mathbf{x}_{-k}))$ is a strictly concave K -person game [38]. Accordingly, the existence of a Nash equilibrium directly results from Theorem 1 and the uniqueness of this equilibrium from Theorem 2 of [39]. This unique Nash equilibrium of considered game is the optimal solution of global minimization problem (11) according to Theorem 2 of [34]. So, the actions determined by Algorithm 2 are feasible and strictly individually rational payoff. For any payoff function which satisfy condition P, the necessary and sufficient condition on the network for a Nash Folk theorem to hold is satisfying condition N (see proof of Theorem 2.12 in [36]). Our payoff function $P(x_k, x_{-k})$ satisfy condition P ($J_k(\mathbf{x}_k, \mathbf{x}_{-k})$ is strictly monotone with respect to its arguments [31]). Therefore, by considering a topology with satisfy condition N, our framework will be converging to a full Nash Folk theorem, i.e. conditions under which all feasible, strictly individually rational payoffs are Nash equilibrium payoffs in a repeated game with patient players.

REFERENCES

- [1] A. Cooper and K. Schefter, "Plug-in electric vehicle sales forecast through 2025 and the charging infrastructure required," *Institute for Electric Innovation (IEI)- Edison Electric Institute (EEI)*, June 2017. [Online]. Available: [http://www.edisonfoundation.net/iei/publications/Documents/IEI_EEI%20PEV%20Sales%20and%20Infrastructure%20thru%202025_FINAL%20\(2\).pdf](http://www.edisonfoundation.net/iei/publications/Documents/IEI_EEI%20PEV%20Sales%20and%20Infrastructure%20thru%202025_FINAL%20(2).pdf)
- [2] K. Clement-Nyns, E. Haesen, and J. Driesen, "The impact of charging plug-in hybrid electric vehicles on a residential distribution grid," *IEEE Transactions on power systems*, vol. 25, no. 1, pp. 371–380, 2010.
- [3] E. Sortomme, M. M. Hindi, S. J. MacPherson, and S. Venkata, "Co-ordinated charging of plug-in hybrid electric vehicles to minimize distribution system losses," *IEEE transactions on smart grid*, vol. 2, no. 1, pp. 198–205, 2011.
- [4] Y. He, B. Venkatesh, and L. Guan, "Optimal scheduling for charging and discharging of electric vehicles," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1095–1105, Sept 2012.
- [5] C. Wu, H. Mohsenian-Rad, and J. Huang, "Vehicle-to-aggregator interaction game," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 434–442, March 2012.
- [6] L. Gan, U. Topcu, and S. H. Low, "Optimal decentralized protocol for electric vehicle charging," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 940–951, May 2013.
- [7] W. J. Ma, V. Gupta, and U. Topcu, "Distributed charging control of electric vehicles using regret minimization," in *53rd IEEE Conference on Decision and Control*, Dec 2016, pp. 4917–4923.
- [8] J. Li, C. Li, Y. Xu, Z. Dong, K. Wong, and T. Huang, "Noncooperative game-based distributed charging control for plug-in electric vehicles in distribution networks," *IEEE Transactions on Industrial Informatics*, 2016.
- [9] Z. Liu, Q. Wu, S. Huang, L. Wang, M. Shahidehpour, and Y. Xue, "Optimal day-ahead charging scheduling of electric vehicles through an aggregative game model," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–1, 2017.
- [10] S. Liu and A. H. Etemadi, "A dynamic stochastic optimization for recharging plug-in electric vehicles," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–1, 2017.
- [11] W. Tang and Y. J. A. Zhang, "A model predictive control approach for low-complexity electric vehicle charging scheduling: optimality and scalability," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1050–1063, 2017.
- [12] E. Yao, V. W. S. Wong, and R. Schober, "Robust frequency regulation capacity scheduling algorithm for electric vehicles," *IEEE Transactions on Smart Grid*, vol. 8, no. 2, pp. 984–997, March 2017.

- [13] T. Zhao and Z. Ding, "Distributed initialization-free cost-optimal charging control of plug-in electric vehicles for demand management," *IEEE Transactions on Industrial Informatics*, vol. PP, no. 99, pp. 1–1, 2017.
- [14] R. Wang, G. Xiao, and P. Wang, "Hybrid centralized-decentralized (hcd) charging control of electric vehicles," *IEEE Transactions on Vehicular Technology*, 2017.
- [15] Q. Huang, Q. S. Jia, Z. Qiu, X. Guan, and G. Deconinck, "Matching ev charging load with uncertain wind: A simulation-based policy improvement approach," *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1425–1433, May 2015.
- [16] A. Brooks, E. Lu, D. Reicher, C. Spirakis, and B. Weihl, "Demand dispatch," *IEEE Power and Energy Magazine*, vol. 8, no. 3, pp. 20–29, May 2010.
- [17] C. Sun, F. Sun, and S. J. Moura, "Nonlinear predictive energy management of residential buildings with photovoltaics & batteries," *Journal of Power Sources*, vol. 325, pp. 723–731, 2016.
- [18] M. Latifi, A. Khalili, A. Rastegarnia, and S. Sanei, "Fully distributed demand response using adaptive diffusion-stackelberg algorithm," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2291 – 2301, Oct 2017.
- [19] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, "A survey on demand response programs in smart grids: Pricing methods and optimization algorithms," *IEEE Communications Surveys Tutorials*, vol. 17, no. 1, pp. 152–178, Firstquarter 2015.
- [20] R. Deng, Z. Yang, M. Y. Chow, and J. Chen, "A survey on demand response in smart grids: Mathematical models and approaches," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 3, pp. 570–582, June 2015.
- [21] J. Forman, J. Stein, and H. Fathy, "Optimization of dynamic battery parameter characterization experiments via differential evolution," in *American Control Conference (ACC)*, 2013. IEEE, 2013, pp. 867–874.
- [22] Z. Ma, S. Zou, L. Ran, X. Shi, and I. A. Hiskens, "Efficient decentralized coordination of large-scale plug-in electric vehicle charging," *Automatica*, vol. 69, pp. 35–47, 2016.
- [23] Z. Ma, S. Zou, and X. Liu, "A distributed charging coordination for large-scale plug-in electric vehicles considering battery degradation cost," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 5, pp. 2044–2052, 2015.
- [24] S. Bahrani and M. H. Amini, "A decentralized framework for real-time energy trading in distribution networks with load and generation uncertainty," *arXiv preprint arXiv:1705.02575*, 2017.
- [25] J. M. Bernardo and A. F. Smith, "Bayesian theory," 2001.
- [26] J.-L. Gauvain and C.-H. Lee, "Maximum a posteriori estimation for multivariate gaussian mixture observations of markov chains," *IEEE transactions on speech and audio processing*, vol. 2, no. 2, pp. 291–298, 1994.
- [27] R. T. Rockafellar, S. Uryasev *et al.*, "Optimization of conditional value-at-risk," *Journal of risk*, vol. 2, pp. 21–42, 2000.
- [28] R. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1443 – 1471, 2002. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378426602002716>
- [29] S. Bahrani and M. H. Amini, "A decentralized trading algorithm for an electricity market with generation uncertainty," *arXiv preprint arXiv:1705.02577*, 2017.
- [30] P. Tabuada, "Event-triggered real-time scheduling of stabilizing control tasks," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1680–1685, 2007.
- [31] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge University Press, 2004.
- [32] G. Loomes and R. Sugden, "Regret theory: An alternative theory of rational choice under uncertainty," *The economic journal*, vol. 92, no. 368, pp. 805–824, 1982.
- [33] A. Rakhlina and K. Sridharan, "Online learning with predictable sequences," *Journal of Machine Learning Research*, vol. 30, 08 2012.
- [34] A. H. Mohsenian-Rad, V. W. S. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid," *IEEE Transactions on Smart Grid*, vol. 1, no. 3, pp. 320–331, Dec 2010.
- [35] D. Fudenberg and E. Maskin, "The folk theorem in repeated games with discounting or with incomplete information," *Econometrica: Journal of the Econometric Society*, pp. 533–554, 1986.
- [36] M. Laclau, "A folk theorem for repeated games played on a network," *Games and Economic Behavior*, vol. 76, no. 2, pp. 711 – 737, 2012.
- [37] *Pennsylvania-New Jersey-Maryland Interconnection (PJM) electricity market*, 2017. [Online]. Available: <http://www.pjm.com/markets-and-operations/energy/day-ahead/Impda.aspx>
- [38] M. J. Osborne and A. Rubinstein, *A course in game theory*. MIT press, 1994.
- [39] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica: Journal of the Econometric Society*, pp. 520–534, 1965.



Milad Latifi received the M.Sc degree in communication engineering from Malayer University, Hamedan, Iran, in 2017. His research interests include advanced signal processing, adaptive filtering, cooperative learning, multi-agent networking, and adaptive optimization. Mr. Latifi is a student member of the IEEE.



Amir Rastegarnia completed his PhD degree in the electrical engineering at the University of Tabriz, Tabriz, Iran, in 2011. In 2011, he joined the Department of Electrical Engineering, Malayer University, as Assistant Professor. His current research interests are theory and methods for adaptive and statistical signal processing, distributed adaptive estimation, as well as signal processing for communications. He is a Member of IEEE.



Azam Khalili received the PhD degree in electrical engineering from the University of Tabriz, Tabriz, Iran, in 2011. In 2011, she joined the Department of Electrical Engineering, Malayer University, as Assistant Professor. Her current research interests are theory and methods for adaptive filtering, distributed adaptive estimation, as well as signal processing for communications. She is a Member of IEEE.



Saeid Sanei (SM05) received his PhD in signal processing from Imperial College London, UK. He has been a member of academic staff in Iran, Singapore, and the UK. He has published three monographs, a number of book chapters, and over 320 papers in peer reviewed journals and conference proceedings. His research interest is in adaptive filtering, cooperative learning, multi-way, multimodal, and multichannel signal processing with applications to biomedical, audio, biometrics, and communication signals and images. He has served as an Associate Editor for the IEEE Signal Processing Letters, IEEE Signal Processing Magazine, and Journal of Computational Intelligence and Neuroscience.