A Bayesian Real-Time Electric Vehicle Charging Strategy for Mitigating Renewable Energy Fluctuations

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

Abstract—A novel pricing and scheduling mechanism is proposed here for Plug-in electric vehicles (PEVs) charging/discharging to track and synchronize with a renewable power generation pattern. Moreover, the proposed mechanism can be used in the demand-side management and ancillary service applications, respectively for the peak shaving and frequency regulation responding. We design a fully distributed stochastic optimization mechanism using Bayesian pure strategic repeated game by which the PEVs optimally schedule their demands. We also use a mixed Bayesian-diffusion Kalman filtering strategy for the customers to collaboratively estimate and track the stochastic price and regulation signals for the upcoming scheduling window. In the proposed paper all the characteristics of the PEVs, as well as the uncertainty about their deriving patterns are considered. As our framework converges to an equilibrium even with incomplete information, is agent-based, and the agents share the information only with their optional neighbors, it is scale-free, robust, and secure.

Index Terms—Bayesian game scheduling, diffusion Kalman filtering, plug-in electric vehicles, renewable energy resources.

I. INTRODUCTION

Today, there is an extensive public attention all over the world for electrifying the transportation sector and integration of the renewable energy resources (RERs) into the power grid. In 2017, the global PEV sales until September were over 764 000 units (46% higher than that in the same period of 2016) [1]. However, increasing the penetration of the RERs and PEVs into the traditional power systems is a critical and can put the reliability and stability of the power grid into danger and degrade its performance and efficiency. This is due to high intermittent nature of the RERs, their inconsistency with energy usage, and high/uncertain power demand of the PEVs. So, it is essential to provide an effective energy management framework to schedule and control the time/amount of charge and discharge of the PEVs and match them with the renewable time/amount power generation.

A. Related Work

Integration of the PEVs to a power system with high wind power penetrations and the application of considering the PEVs as a regulation services provider has been addressed in [2]. The western Danish power system where the total share of annual wind power production is more than 27% of the electrical energy demand was chosen for study in this analysis. Jin *et al.* in [3] formulated the energy consumption scheduling (ECS) problem as a stochastic optimization problem based on a queuing model which uses Lyapunov dynamic optimization technique to minimize the time average cost of using non-renewable energy sources. Zhang *et al.* [4] considered a delay-optimal ECS of the PEVs with queue mapping to convert the PEV queue to the charge demand queue at a charging station equipped with renewable energy generation devices and multiple charge points. Further, an independent Markov decision process (MDP) has been described in this investigation to take into account uncertainty of the PEV arrival, the intermittence of the RES, and the variation of the grid power price. Chen and Duan [5] proposed a method for optimal integration of the PEVs in the microgrids dealing which uncertainties associated with the daily distance driven of the PEVs, load values, and electricity market price and considered an optimal number of parking numbers under optimal scheduling of the PEVs. The problem of matching the stochastic wind power with the PEVs charging schedule considering the optimality of the expected charging cost has been studied by Huang *et al.* [6]. This stochastic multi-stage matching problem has been formulated as an MDP by aggregating the PEVs according to their remaining parking time, and was solved by simulation-based policy improvement method. A real-time load elasticity tracking and pricing for the ECS problem using predicting PEV charging behavior of the consumers at different prices has been investigated by Soltani *et al.* [7]. The dependencies on price responsiveness among consumers are captured by a conditional random field (CRF) model in this work. Lee *et al.* [8] perused the price competition among the PEV charging stations equipped with renewable power generators where each station sets its electricity price to maximize its revenue by competing with its neighbors. The proposed model in this work is a game theoretic model in which some relevant physical constraints such as the transmission line capacity, the distance between the PEV and charging station, and the number of charging outlets at the charging station are taken into account.

A two layer hierarchical control scheme for the coordination of PEVs and wind turbines in a microgrid has been modeled in [9]. The main feature of this scheme is that it incorporates the non-Gaussian uncertainty and partially dispatchability of wind power, as well as the PEV uncertainty. A two-stage (day-ahead/real-time) mechanism was established in [10] for integrating the massive PEVs into the power grid, tackling the RESs' uncertainties, and coordinating them to reduce the energy cost and peak to average ratio (PAR) of the system. At the first stage, based on the prediction of future energy requests and considering the elastic charging property of the PEVs, an offline day-ahead optimal energy generation scheduling problem was formulated to determine the energy generation. Then at the second stage, based on the planned energy generation, an adaptive real-time charging strategy is developed to determine the charging rate of each vehicle in a dynamic manner. A charging rate compression algorithm which tremendously reduces the complexity of the problem solving has been also developed in this paper. In [11], the ECS problem has been formulated using MDP and solved using a distributed simulation-based policy improvement. The proposed method takes into account the uncertain wind power sources and can be improved from heuristic and experience based policies. A case study in Beijing (China) was developed to explore the potential of wind power of high-rise buildings in supporting the PEVs charging. Yang *et at.* [12] considered a grid-connected microgrid model which consists of a logistics distribution system, where the PEVs depart from the depot, deliver the goods to multiple demand loads, and then return to the depot.

A coordinated dispatch strategies of the PEVs to smooth renewable energy and load fluctuations of the microgrid has been developed, while ensuring the quality of logistics services. They established an ECS method using a self-adaptive imperialist competitive algorithm to optimize the driving routes, fast-charging time and regular-charging/discharging strategies of multiple PEVs. Luo *et al.* [13] studied the problem of stochastic dynamic pricing and energy management policy for the PEV charging service providers in the presence of the energy storage system and multiple uncertainty sources. They offered guidelines for charging service providers to the determine proper charging prices and manage the electricity to balance the competing objectives of improving profitability, enhancing customer satisfaction, and reducing impact on power grid in spite of these uncertainties. Moreover, a new metric to assess the impact on power grid, without completely solving the power flow equations has been proposed in this paper.

B. Our Contributions

Based on the drawbacks in previous works, we are motivated to contribute to:

Designing a robust and scalable game mechanism: The availability of some forms of information to all the customers and to the utility company is fundamental to coordinate their actions to result in a socially fair solution. However, this raises some challenges such as requiring a two-way data communication network with high communication/computation capability,

preserving the customers' privacy, and ensuring the power grid security and reliability. As in our framework the customers communicate and share information only with their optional neighbors and the proposed Bayesian game converges to an equilibrium even with incomplete information, the mechanism is scale-free, requires low communication/computation, and is more secure and reliable. Further, as long as the topology and information sharing rule of the communication network satisfy some conditions as described in Section IV of this paper, the power grid is robust to noise, link failures, and incomplete information.

Providing a universal real-time ECS: As there are several uncertainty sources (market price, power generation, power consumption, etc.) in the power grid, the time-ahead ECS methods usually lead to sub-optimal or even non-optimal solutions. So, we develop a real-time mechanism using the rolling horizon technique and online adaptive Bayesian-diffusion strategy to optimize the solution. The proposed method is universal in the sense that it can be used for both demand-side management (DSM) and ancillary service necessities.

Considering the customer's attribution: The works in the literature are mostly focusing on using the flexibility of the battery of the PEVs for DSM and frequency regulation objectives without explicitly addressing the PEV's energy demand limitation at the end of each scheduling window. We proceed in this matter by formulating a pure strategic game and involving the PEV's characteristics in detail. We develop a binary Markov chain model updated using Bayesian method to predict the PEVs plug-in and departure times. Moreover, we provide a robust model by which the total energy demand of each PEV for its next trip is estimated.

Mitigating the adverse effect of the PEV and RER: By designing a novel real-time behavioral price policy, the customers in the considered model try to match their PEVs' charging/discharging pattern with the stochastic highly fluctuating power generated from the RERs by tracking its variation. This results in making the balance between supply and demand sides and jointly smoothing out the total power generation and consumption curves.

The remainder of this paper is outlined as follows. We demonstrate the system model in Section II. Section III introduces the agent-based game theoretic PEV scheduling problem. The Bayesian-diffusion-Kalman filtering solution is developed in Section IV. The simulation results are presented and assessed in Section V, and Section VI provides the concluding remarks.

Notation: lower- (upper-) case boldface letters denote column vectors (matrices), and calligraphic letters stand for sets. The notation $\mathbb{E}[\cdot]$ denotes expectation operator, $|\cdot|$ denotes the cardinality operation of a set, $\mathrm{Tr}(\cdot)$ denotes the matrix trace operation, $\mathrm{Pr}(\cdot)$ denotes the probability operation, $\times_{n=1}^{N}$ denotes the Cartesian product over N elements, $\otimes_{n=1}^{N}$ denotes the σ-algebra [14] product on the Cartesian product, and $\tilde{a}$ implies that a is a random variable. We denote the σ-algebra of Borel [15] subsets of a by $\mathscr{B}(a)$ and its transpose by $a^{\top}$. Further, $\mathcal{A}_1 \Delta \mathcal{A}_2 = (\mathcal{A}_1/\mathcal{A}_2) \bigcup (\mathcal{A}_2/\mathcal{A}_1)$ and $\mathcal{N}_n/j$ is the set of player n's neighbors including itself and excluding neighbor j. The subscript a^t stands for value of variable a for a period

t, while, to denote its value at a time slot h within period t we use a^h instead of $a^{h,t}$ for simplicity.

II. SYSTEM MODEL

We consider a smart micro-grid system (network) comprising of one utility company which procures energy from the wholesale market and its own renewable farm, and serves a set $\mathcal{N} \triangleq \{1, \cdots, N\}$ of $N = |\mathcal{N}|$ PEVs of residential/commercial/industrial owners (called the customers) as in Fig. 1. Each customer's site is equipped with an energy consumption manager (ECM) device responsible for scheduling of its PEV power consumption, i.e., determining the optimal PEV charging/discharging rate/time. The customers participate in the ECS program for a long time divided into set $\mathcal{T} \triangleq \{1, \cdots, T\}$ of equal length periods, while their actions have a periodic pattern within each $t \in \mathcal{T}$ (e.g., 1 year with $|\mathcal{T}| = T = 365$ and each period $t \in \mathcal{T}$ as one day). We divide each period $t \in \mathcal{T}$ into a set $\mathcal{H}^t$ of H^t equal time slots (e.g., one day as $|\mathcal{H}^t| = H^t = 96$ with a slot duration of 15 minutes). It is assumed that each ECM[1] $n \in \mathcal{N}$ is connected to the utility company and its neighbors[2] $m \in \mathcal{N}_n \subseteq \mathcal{N}$ with some proper two-way communication protocol. At each period t, each customer n tries to cooperatively manage its PEV energy consumption throughout its feasible scheduling window $\mathcal{H}_n^t \triangleq \{\alpha_n^t, \cdots, \beta_n^t\}$ (with $|\mathcal{H}_n^t| = H_n^t = \beta_n^t - \alpha_n^t + 1$), where α_n^t is the PEV plug-in time and β_n^t is the last time slot at which the PEV had to reach to a desirable charge level $E_n^{d,t}$. Further, we denote the remained time slots from slot τ to the end of scheduling window $\mathcal{H}_n^t$ at period t with $\mathcal{H}_n^\tau \subseteq \mathcal{H}_n^t$. We consider three modes for the PEV status, namely, *off* mode when the PEV is unplugged or finished its task, *asleep* mode when is plugged-in but does not charge/discharge, and *awake* mode when is charging/discharging. For PEV n, we define its operation state at time slot $h \in \mathcal{H}_n^t$ in period t as a couple $s_n^h(r_n^h, d_n^h)$, where r_n^h is the number of remaining time slots to complete the current task and d_n^h is the number of time slots for which the current task can be delayed. As an insight, the state of PEV n in period t throughout its scheduling window $|\mathcal{H}_n^t| = 9$ hours with $r_n^h = 4$ and $d_n^h = H_n^t - r_n^h = 5$ is denoted in Fig. 1. As is depicted, the PEV is in off mode when is unplugged (out of scheduling window $\mathcal{H}_n^t$) or when has finished its work, and does not consume more power.

A. PEV characterization

Let us denote PEV n operation mode in time slot $h \in \mathcal{H}_n^t$ with binary decision variables $x_{n,c}^h \in \{0 \text{ (asleep)}, +1 \text{ (awake-charge)}\}$ and $x_{n,d}^h \in \{0 \text{ (asleep)}, -1 \text{ (awake-discharge)}\}$. In an optimal
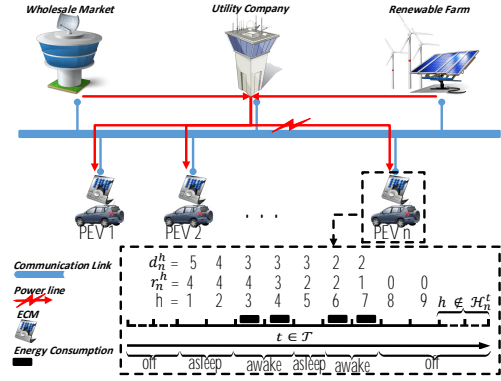

Fig. 1. The considered power grid block-diagram model.

scheduling of the PEV n's operation at each t, the ECM faces the following constraints:

$$
\begin{aligned}
&\sum_{h=\tau \geq \alpha_n^t}^{\beta_n^t} \left(\eta_{n,c} \cdot x_{n,c}^h + \frac{x_{n,d}^h}{\eta_{n,d}} \right) \cdot x_n^{rat} = E_n^{d,t} - (E_n^{\tau-1} + E_n^0), \\
&E_n^h = E_n^{h-1} + \left(\eta_{n,c} \cdot x_{n,c}^h + \frac{x_{n,d}^h}{\eta_{n,d}} \right) \cdot x_n^{rat}, \\
&E_n^{min} \leq E_n^h \leq B_n^{cap}, \ \forall h \in \mathcal{H}_n, \ E_n^{lb} \leq E_n^{d,t} \leq E_n^{ub}, \\
&x_{n,c}^h \cdot x_{n,d}^h = 0, \ \forall h \in \mathcal{H}_n, \\
&x_{n,c}^h + x_{n,d}^h = 0, \ \forall h \notin \mathcal{H}_n
\end{aligned}
$$

$$\text{(1)}$$

where E_n^h, E_n^0, E_n^{min}, and B_n^{cap} are the energy level at the end of slot h, initial energy level, minimum acceptable energy level, and the battery capacity, respectively. Coefficients $\eta_{n,c}, \eta_{n,d} \in (0, 1]$ denote charging and discharging efficiency and x_n^{rat} is the rated power at which the PEV is charged/discharged. The first dynamic term of (1) is for providing the required energy level for the next trip before the departure time of the PEV. The second line denotes the evolution of the energy level of the battery of the PEV. The constraints in the third line bound the energy level and denote the tolerable deviation from $E_n^{d,t}$, respectively. The fourth line of equation (1) implies that the PEV cannot be charged and discharged at the same time and the fifth line ensures that the PEV cannot be charged/discharged when it is unplugged. Each ECM $n \in \mathcal{N}$ can infer the posterior state $s_n^{h+1}(r_n^h - x_{n,c}^{h+1} + x_{n,d}^{h+1}, d_n^h + x_{n,c}^{h+1} - x_{n,d}^{h+1} - 1)$ of its PEV in the next time slot $h+1$ from the prior state $s_n^h(r_n^h, d_n^h)$ and the current action $\boldsymbol{x}_n^h = [x_{n,c}^h \cdot x_n^{rat}, \ x_{n,d}^h \cdot x_n^{rat}]^\top$. Accordingly, the ECM updates its feasible set $\mathcal{X}_n^h := \{\boldsymbol{x}_n^h \mid \text{constraints in (1)}\}$ at each $h \in \mathcal{H}^t$ and resets them at each new $t \in \mathcal{T}$.

B. Energy source model

We assume that the renewable power is produced by a wind farm comprising of a set $\mathcal{K} \triangleq \{1, \cdots, K\}$ of $|\mathcal{K}| = K$ turbines. The power $\tilde{P}_{k,r}^h$ drawn from each turbine $k \in \mathcal{K}$

---

[1]Throughout the paper, we use ECM, customer, and PEV interchangeably.
[2]Set of customers who customer n can or would like to interact with them, including itself is called neighborhood of n and denoted by $\mathcal{N}_n$.

at each slot $h \in \mathcal{H}^t$ is calculated as [16]:

$$\tilde{P}_{k,r}^h := \begin{cases} \frac{1}{4}\pi R_k^2 \varpi_k^h (v_k^h)^3 \left(1 + \frac{v_{k,d}^h}{v_k^h}\right) \left(1 - \left(\frac{v_{k,d}^h}{v_k^h}\right)^2\right), & v_k^{in} \leq v_k^h \leq v_k^{rat} \\ P_k^{rat}, & v_k^{rat} < v_k^h \leq v_k^{out} \\ 0, & v_k^h < v_k^{in} \text{ or } v_k^h > v_k^{out} \end{cases} \tag{2}$$

where R_k [m] is the rotor radius of turbine k, and ϖ_k^h [kg/m^3], v_k^h [m/s], and $v_{k,d}^h$ [m/s] are air density, wind speed, and downstream wind speed at the site of turbine k, respectively. Further, P_k^{rat} [MW] is the rated power output of wind turbine k, and v_k^{in}, v_k^{out}, and v_k^{rat} [m/s] are cut-in, cut-out, and rated wind speeds of wind turbine k, respectively. The variation in wind velocity v_k^h is best to follow by the Weibull probability distribution function (PDF) with shape ω_k^h and scale c_k^h parameters. So, the probability of wind speed v_k^h at the site of turbine k during the time slot h is given by [17]:

$$\pi_w^t(v_k^h) = \frac{\omega_k^h}{c_k^h} \left(\frac{v_k^h}{c_k^h}\right)^{\omega_k^h - 1} e^{-\left(\frac{v_k^h}{c_k^h}\right)^{\omega_k^h}} \tag{3}$$

where, mostly, the wind speed is considered to have Weibull distribution with $\omega_k^h = 2$, which is specifically known as the Rayleigh distribution. Moreover, the higher the c_k^h is, the greater the number of days/hours that have high winds (see [18] for details). The utility company needs to know the total power drawn from the renewable farm $\sum_{k \in \mathcal{K}} \tilde{P}_{k,r}^h$ and the aggregate non-PEV load demand $\sum_{n \in \mathcal{N}} l_n^h$ for each time slot $h \in \mathcal{H}^t$ before the scheduling window. Accordingly, the utility company makes a contract with the wholesale market for buying the total base-power P_b^h required for making the balance between the supply and demand sides. However, there is always inaccuracy in determining P_b^h ahead of time which must be compensated in real-time by some regulation power $\tilde{w}^h$. The required regulation power $\tilde{w}^h$ to make the real-time balance is calculated as follows:

$$\tilde{w}^h = \left(P_b^h + \sum_{k \in \mathcal{K}} \tilde{P}_{k,r}^h\right) - \sum_{n \in \mathcal{N}} l_n^h \tag{4}$$

It is rational to assume that the utility company determines P_b^h properly at each slot so that $\tilde{w}^h$ follows a normal distribution $\mathcal{N}(0, \sigma_r^2)$[3] as:

$$\pi_r^t(w^h) = \frac{1}{\sqrt{2\pi\sigma_r^2}} e^{-\frac{(w^h)^2}{2\sigma_r^2}} \tag{5}$$

C. Price design

In the utility company view, the most suitable tool to encourage the customers for following a specific energy consumption pattern is establishing an effective price policy. As in the pricing policy the customers' payment depends on the time and amount of consumed energy, the customers try to consume more energy at low-price times and vice versa. So, it is

---

[3]The PDF of $\tilde{w}^h$ is assumed to be Gaussian $\mathcal{N}(m, \sigma_r^2)$ when the PEV penetration to the power grid is large enough to support some base loads as well as tracking the wind power production.

---

rational to assume that the utility company determines its price parameter at slot $h \in \mathcal{H}^t$ in period t based on:

$$\tilde{p}^h = \tilde{p}_b^h + \gamma^h \tilde{p}_{sh}^h, \quad \tilde{p}_{sh}^h = \left(\sum_{n \in \mathcal{N}} (x_{n,c}^h + x_{n,d}^h) - \tilde{w}^h\right)^2 \tag{6}$$

where random variable $\tilde{p}_b^h$ is the real-time baseline price which is unknown to the customers a priori. When total non-PEV load demand $\sum_{n \in \mathcal{N}} l_n^h$ is greater (lower) than total procured power $P_b^h + \sum_{k \in \mathcal{K}} \tilde{P}_{k,r}^h$, e.g., $\tilde{w}_t^h < 0 \ (> 0)$, the customers are encouraged to discharge (charge) their PEVs to reduce their payments through minimizing $\tilde{p}_{sh}^h$ (called the shadow price). This procedure leads to track random variable $\tilde{w}^h$ which results in the power balance.

III. PROBLEM FORMULATION

A. Sources of Uncertainty

The deriving patterns of the PEVs are stochastic in nature. So, to provide an efficient charging/discharging schedule at each period t, the plug-in time (i.e., α_n^t), the plug-out time (i.e., β_n^t), and the total energy demand (i.e., $E_n^{d,t}$) of each PEV n during the connection time $\mathcal{H}_n^t$ need to be estimated accurately. To estimate the stochastic parameters α_n^t and β_n^t, one can model the available time (i.e., the length of $\mathcal{H}_n^t$) of PEV n as a binary Markov chain with state = 1 denoting that the PEV is available to the ECM (e.g., is in the parking lot) and state = 0 denoting that the PEV is unavailable. Staring from an initial state, we can infer the feasible scheduling window $\mathcal{H}_n^t$ using the transition probabilities of the corresponding Markov chain.

Let's denote the status of PEV n within the period t by the homogenous and stationary Markov chain $\mathcal{M}_n^t = \{M_n^0, \cdots, M_n^H\}$ with transition probabilities $\rho_{n,ij} = \Pr(M_n^{h+1} = j | M_n^h = i)$, $i, j \in \{0, 1\}$. The probability of the length of time at which the PEV may be available (i.e., remains in state 1) is computed by $\rho_n^t = \rho_{n,01}/(\rho_{n,01} + \rho_{n,10})$, represents the long-run proportion of time when the Markov chain is in state 1 [19]. Letting $\mathcal{M}_n^t = \{m_n^0, \cdots, m_n^H\}$ denote a fully observed realization of $\mathcal{M}_n^t$, conditionally to $M_n^0 = 1$, the distribution of the observed sequence is then [20]:

$$f(M_n^t | m_n^0, \zeta_n^t) = (1 - \rho_{n,01})^{h_{00}} \rho_{n,01}^{h_{01}} \rho_{n,10}^{h_{10}} (1 - \rho_{n,10})^{h_{11}} \tag{7}$$

where h_{ij} is the number of one-step transition from state i to state j until time-slot h and the unknown probability parameter $\zeta_n^t = (\rho_{n,01}, \rho_{n,10}) \in [0, 1]$ can be inferred using a proper estimation technique, such as the maximum likelihood estimates. However, as we aim to provide a real-time adaptive PEV charging strategy, Bayesian inference, which incorporate a certain type of dependence between the components of the parameters in diferent periods $t \in \mathcal{T}$, seems to be the most suitable choice.

Proposition 1. *Considering Jeffreys' prior distribution $\pi(\zeta_n^t)$ on parameters $\rho_{n,01}$ and $\rho_{n,10}$, the conditional posterior*

density becomes:

$$\pi(\zeta_n^t|\mathcal{M}_n^t) \propto \left[h(\rho_{n,01}+\rho_{n,10}) - 1 + (1-\rho_{n,01}-\rho_{n,10})^h\right]^{1/2} \times$$
$$\left[h\rho_{n,01}(\rho_{n,01}+\rho_{n,10}) + \rho_{n,10}(1-\rho_{n,01}-\rho_{n,10})^h\right]^{1/2} \times \rho_{n,01}^{h_{01}-1/2}$$
$$(1-\rho_{n,01})^{h_{00}-1/2}\rho_{n,10}^{h_{10}}(1-\rho_{n,10})^{h_{11}-1/2}(\rho_{n,01}+\rho_{n,10})^{-2} \tag{8}$$

Proof. See Section 2 of [21]. □

Accordingly, the Bayesian estimator becomes $\mathbb{E}[\zeta_n^t|\mathcal{M}_n^t] = \int \zeta_n^t \pi(\zeta_n^t|\mathcal{M}_n^t)d\zeta_n^t$, which its integral is difficult to calculate in practice. One can approximate the Bayesian estimator by means of Markov chain Monte Carlo techniques [22].

To infer $E_n^{d,t}$, we propose the use of a deterministic set to model the uncertainty, which results in a robust decision about the desired energy demand at each period t. Inspired by the work in [23], we consider the following uncertainty set of the desired energy demand at each time period t in the planning horizon $\mathcal{T}$:

$$\mathcal{E}_n^t(\boldsymbol{E}_n^{d,t}, \acute{\boldsymbol{E}}_n^{d,t}, \Delta_n^t) := \left\{ \hat{\boldsymbol{E}}_n^{d,t} \in \mathbb{R}^{N_n} : \sum_{i\in\mathcal{N}_n} \frac{|\hat{E}_i^{d,t}-E_i^{d,t}|}{\acute{E}_i^{d,t}} \le \Delta_n^t \right.$$
$$\left. \hat{E}_i^{d,t} \in [E_i^{d,t}-\acute{E}_i^{d,t}, E_i^{d,t}+\acute{E}_i^{d,t}], \forall i\in\mathcal{N}_n, t\in\mathcal{T} \right\} \tag{9}$$

where $\mathcal{N}_n$ is the set of PEVs managed by the ECM of customer n with cardinality $|\mathcal{N}_n| = N_n$ and $\boldsymbol{E}_n^{d,t} = (E_i^{d,t}, \forall i\in\mathcal{N}_n)$, $\hat{\boldsymbol{E}}_n^{d,t} = (\hat{E}_i^{d,t}, \forall i\in\mathcal{N}_n)$, and $\acute{\boldsymbol{E}}_n^{d,t} = |\boldsymbol{E}_n^{d,t} - \hat{\boldsymbol{E}}_n^{d,t}|$ are the true vector, the estimated vector, and the deviation vector of the uncertain energy demands for the PEVs under control of customer n, respectively. Further, the parameter Δ_n^t is the budget of uncertainty (confidence level parameter), taking values between 0 and N_n, and determined by the customer according to its attribution. As Δ_n^t increases from 0 to N_n, the size of the uncertainty set enlarges, and the resulting robust decision is more conservative and the system is protected against a higher degree of uncertainty.

B. The Optimization Problem

Assuming the grid's customers are selfish, their participation in any ECS program is to minimize their payments. In this way, the social objective function of the customers is defined as follows:

$$\min_{\boldsymbol{X}_1^t,\cdots,\boldsymbol{X}_N^t} \sum_{t\in\mathcal{T}}\sum_{n\in\mathcal{N}}\sum_{h\in\mathcal{H}} \mathbb{E}[\tilde{p}^h] \cdot \left(l_n^h + (x_{n,c}^h + x_{n,d}^h)\cdot x_n^{rat}\right)$$
$$\text{s.t. } \boldsymbol{X}_n^t \in \mathcal{X}_n^h, \forall n\in\mathcal{N}, h\in\mathcal{H}^t, \hat{\boldsymbol{E}}_n^{d,t}\in\mathcal{E}_n^t, \text{ and } t\in\mathcal{T} \tag{10}$$

where $\boldsymbol{X}_n^t \triangleq [\boldsymbol{x}_n^1,\cdots,\boldsymbol{x}_n^H]^\top$ is the operation matrix of PEV n throughout the scheduling window $\mathcal{H}^t$. Further, $\mathcal{X}_n^h$ is the feasible set updated at slot h in period t comprises of the set of profiles satisfying constraints in (1). Centrally Solving problem (10) imposes a very high communication/computation burden on the central controller, put into danger the privacy of the customer as their information must be sent to the central controller, and is not reliable as any failure in the central controller would take down the whole system. Although the semi-decentralized solutions (such as the classical Nash game method) can alleviate some drawbacks of the centralized

solutions, however, they still need each customer to send some parts of its information to all other customers. So, these methods are not fully scalable, reliable, or secure. Due to the drawbacks of centralized and semi-decentralized solutions, problem (10) must be solved in a fully distributed manner with only local cooperation and partially available information. However, this problem is stochastic and spatially-temporally coupled (due to the price parameters $\tilde{p}^h$ and constraints in (1)) which complicates providing an optimal distributed solution. To solve this problem, we suggest developing a ***stochastic Bayesian game with private and incomplete information*** mechanism defined as $\mathscr{G} = \left(\mathcal{X}_n^h, (\mathcal{S}_n^h, \mathscr{S}_n^h), U_n^t, \pi_n^t\right)_{n=1}^N$ with the following components [24]:

Players: Set $\mathcal{N}$ of a finite N number of PEVs.

Action space $\mathcal{X}_n^h$: Is a nonempty compact set (metric space) of feasible operation modes over strategy $\boldsymbol{X}_n^t$ of PEV n.

Information space $(\mathcal{S}_n^h, \mathscr{S}_n^h)$: Is a measurable space defined on $\mathcal{X}_n^h$, where $\mathcal{S}_n^h$ is player n's nonempty type set with state $(s_n^h(r_n^h, d_n^h), \forall h\in\mathcal{H}^t)$, and measurable space $\mathscr{S}_n^h$ is σ-algebras on $\mathcal{S}_n^h$.

Payoffs U_n^t: Represents player n's objective function $U_n^t : \mathcal{S}^t \times \mathcal{X}^t \to \mathbb{R}$ to be maximized, defined on the set of action spaces $\mathcal{X}^t =: [\times_{n=1}^N \mathcal{X}_n^h = \mathcal{X}_n^h \times \mathcal{X}_{-n}^h]_{h\in\mathcal{H}^t}$ and bounded/finite types $\mathcal{S}^t =: [\times_{n=1}^N \mathcal{S}_n^h = \mathcal{S}_n^h \times \mathcal{S}_{-n}^h]_{h\in\mathcal{H}^t}$ with $\left([\otimes_{n=1}^N \mathscr{S}_n^h]\otimes[\otimes_{n=1}^N \mathscr{B}(\mathcal{X}_n^h)], \mathscr{B}(\mathbb{R})\right)$-measurable, where $\mathcal{X}_{-n}^h$ and $\mathcal{S}_{-n}^h$ are the set of action and type spaces of all the players except for player n with strategy $\boldsymbol{X}_{-n}^h$ at slot h of period t, respectively. This payoff function is dual in negative direction of problem (10) comminuted corresponding to PEV n throughout one scheduling window as:

$$\max_{\boldsymbol{X}_n^h\in\mathcal{X}_n^h} U_n^t(\boldsymbol{X}_n^h, \boldsymbol{X}_{-n}^h)$$
$$= -\sum_{h\in\mathcal{H}}\left(\mathbb{E}[\tilde{p}_b^h] + \gamma^h\left(\sum_{n\in\mathcal{N}}(x_{n,c}^h + x_{n,d}^h)\cdot x_n^{rat} - \mathbb{E}[\tilde{w}^h]\right)^2\right)$$
$$\cdot \left(l_n^h + (x_{n,c}^h + x_{n,d}^h)\cdot x_n^{rat}\right) \tag{11}$$

Information structure π^t: Is a probability distribution measure on the measurable space $[(\otimes_{n=1}^N \mathcal{S}_n^h, \otimes_{n=1}^N \mathscr{S}_n^h)]_{h\in\mathcal{H}^t}$ denoting the common prior over type profiles, with π_n^t the marginal probability measure induced by π^t on $\mathcal{S}_n^h$, i.e., the probability measure on $(\mathcal{S}_n^h, \mathscr{S}_n^h)$ defined by $\pi_n^t(S_n^h) := \pi^t(S_n^h \times \mathcal{S}_{-n}^h)$, for every $S_n^h \in \mathscr{S}_n^h$.

Definition 1. *A pure strategy $\boldsymbol{X}_n^\tau$ for player n in Bayesian stochastic game $\mathscr{G}$ is a $(\mathscr{S}_n^\tau, \mathscr{B}(\mathcal{X}_n^\tau))$-measurable map $\boldsymbol{X}_n^\tau : \mathcal{S}_n^\tau \longmapsto \mathcal{X}_n^\tau$ with the interpretation that, upon learning his type $S_n^\tau(s_n^\tau, \tilde{p}_b^\tau, \tilde{w}^\tau, S_{-n}^\tau) \in \mathcal{S}_n^\tau$, player n selects strategy $\boldsymbol{X}_n^\tau(S_n^\tau)$ from the set $\mathcal{X}_n^\tau$ for the entire scheduling horizon $\mathcal{H}_n^\tau$.*

In deploying the proposed game in real-time, each player in the Bayesian game adopts a behavioral or mixed infinite strategy according to its prior $\pi_n^t(S_n^h)$. The existence proof of a Bayes-Nash equilibrium in this situation is straightforward (e.g., see [25], [26] for detailed analysis and examples). However, satisfying the temporally-coupled energy demand constraint (i.e., the first term of (1)) and determining an optimal schedule need developing a pure strategic game. So, we must convert the behavioral/mixed strategy into a pure

strategy for each sub-scheduling window $\mathcal{H}_n^\tau$ using some purification technique[4]. In the proposed game $\mathcal{G}$, at each slot, the customers update their beliefs according to Bayes theory (14). So, the best choice for the purification is matching the rolling horizon theory [28] with updating procedure to ensure that at any individual condition and preference, the constraints in (1) are not violated and the solution is optimal.

Theorem 1. *The behavioral/mixed strategy of game $\mathcal{G}$ has an equivalent pure strategy, and there exists a pure strategy Nash equilibrium (PSNE) for considered game $\mathcal{G}$ caused by applying the rolling horizon technique on set of pure strategies $\mathbf{X}^h := \times_{n=1}^N \mathbf{X}_n^h$ for all $h \in \mathcal{H}^t$ in each $t \in \mathcal{T}$ so that:*

$$U_n^t(\mathbf{X}_n^{*h}, \mathbf{X}_{-n}^{*h}) \geq U_n^t(\mathbf{X}_n^h, \mathbf{X}_{-n}^{*h})$$

where $\mathbf{X}_n^{*h}$ and $\mathbf{X}_{-n}^{*h}$ denote the equilibrium strategies of player n and all other players decided at slot h for the rest of the slots $H^t - h + 1$, respectively.

Proof. See Appendix A. □

In the next section, we provide a fully distributed Bayesian mechanism for each customer n to learn his type $S_n(s_n^h, \tilde{p}_b^h, \tilde{w}^h, \mathcal{S}_{-n})$ and calculate his equilibrium strategy $\mathbf{X}_n^{*h}(S_n^h)$.

IV. BAYESIAN REAL-TIME SOLUTION

To construct a pure strategy for each customer $n \in \mathcal{N}$ we need to provide an agent based method by which each customer can sequentially estimate the baseline price signal vector $\tilde{\mathbf{p}}_b^h = [\tilde{p}_b^h, \cdots, \tilde{p}_b^H]^\top$ and regulation parameter vector $\tilde{\mathbf{w}}^h = [\tilde{w}^h, \cdots, \tilde{w}^H]^\top$ at each slot h. To do this and to increase the accuracy and robustness of the estimation process, we develop a method in which each customer sequentially updates its belief on the upcoming price and regulation signals and share his information with his neighbors. At each slot $h \in \mathcal{H}^t$, we need to determine and update the PDFs $\pi_{\mathbf{p}}^t$ and $\pi_{\mathbf{w}}^t$ of $\tilde{\mathbf{p}}_b^h$ and $\tilde{\mathbf{w}}^h$ to compute $\hat{\mathbf{p}}_b^h = \mathbb{E}_{\pi_{\mathbf{p}}^t}[\tilde{\mathbf{p}}_b^h]$ and $\hat{\mathbf{w}}^h = \mathbb{E}_{\pi_{\mathbf{w}}^t}[\tilde{\mathbf{w}}^h]$. At the end of each slot h the retailer announces true prices $\tilde{p}_b^h$ and $\gamma^h \tilde{p}_{sh}^h$ of (6), regulation signal $\tilde{w}^h$, and probably some explanatory variables. Accordingly, we model the observation of customer n as $\mathbf{y}_n^h = [y_{n,b}^h,\ y_{n,sh}^h]^\top$ determined by an unknown stochastic parameter $\tilde{\mathbf{\Theta}}^h = [\tilde{\mathbf{p}}_b,\ \tilde{\mathbf{w}}]$, where $y_{n,b}^h$ and $y_{n,sh}^h$ are his expected payment at the end of the scheduling window determined at slot h based on the most recent information about the baseline price, shadow prices, and other explanatory variables. We model any explanatory data (weather condition, fossil fuel price, price in previous days, etc.) available to customer n at slot h by $\mathbf{z}_n^h$. Accordingly, customer n's parameter estimation problem within each period $t \in \mathcal{T}$ at slot h takes the following form:

$$\hat{\mathbf{\Theta}}^h = \underset{\tilde{\mathbf{\Theta}} \in \mathcal{O}}{\arg\min}\, F_n(\tilde{\mathbf{\Theta}}^h, \mathbf{y}_n^h, \mathbf{z}_n^h) \qquad (12)$$

where $\mathcal{O}$ is a compact set which captures possible prior information of slot $h - 1$. Local function measure $F_n(\cdot)$ of

customer n is a convex function of $\tilde{\mathbf{\Theta}}^h$. A popular optimal choice for $F_n(\cdot)$, often advocated in Bayesian theory, is the Kullback-Leibler divergence (KLD) measure. This function is defined as follows:

$$D\big(\hat{\pi}_n^t(\tilde{\mathbf{\Theta}}^h|\cdot)||\pi^t(\tilde{\mathbf{\Theta}}^t|\cdot)\big) = \mathbb{E}_{\hat{\pi}_n^t}\left[\log \frac{\hat{\pi}_n^t(\tilde{\mathbf{\Theta}}^h|\cdot)}{\pi^t(\tilde{\mathbf{\Theta}}^t|\cdot)}\right] \qquad (13)$$

where determining density $\hat{\pi}_n^t(\tilde{\mathbf{\Theta}}^h|\cdot)$ is the goal of customer n. The probability distribution of observation $\mathbf{y}_n^h$ is presented with a density $f_n^t(\mathbf{y}_n^h|\mathbf{z}_n^h, \tilde{\mathbf{\Theta}}^h)$. In the provided sequential Bayesian-diffusion estimation mechanism, each customer $n \in \mathcal{N}$ simultaneously estimates parameter $\tilde{\mathbf{\Theta}}^t$ at slot h by exploiting a prior distribution $\pi_n^t(\tilde{\mathbf{\Theta}}^{h-1}|\mathbf{y}_n^{0:h-1}, \mathbf{z}_n^{0:h-1})$ which quantifies the accumulated knowledge about $\tilde{\mathbf{\Theta}}^t$ from past observations $\mathbf{y}_n^{0:h-1} = \{\mathbf{y}_n^0, \cdots, \mathbf{y}_n^{h-1}\}$ and regressors $\mathbf{z}_n^{0:h-1} = \{\mathbf{z}_n^0, \cdots, \mathbf{z}_n^{h-1}\}$ up to slot h. According to Bayes' theorem, the observation $\mathbf{y}_n^h$ and the regressor $\mathbf{z}_n^h$ revealed at the end of slot h, are sequentially incorporated as:

$$\pi_n^t(\tilde{\mathbf{\Theta}}^h|\mathbf{y}_n^{0:h}, \mathbf{z}_n^{0:h}) \propto f_n^t(\mathbf{y}_n^h|\mathbf{z}_n^h, \tilde{\mathbf{\Theta}}^h)\pi_n^t(\tilde{\mathbf{\Theta}}^{h-1}|\mathbf{y}_n^{0:h-1}, \mathbf{z}_n^{0:h-1}) \qquad (14)$$

Common model to describe $f_n^t(\mathbf{y}_n^h|\mathbf{z}_n^h, \tilde{\mathbf{\Theta}}^h)$ and $\pi_n^t(\tilde{\mathbf{\Theta}}^h|\mathbf{y}_n^{0:h}, \mathbf{z}_n^{0:h})$ is the exponential family of distributions as [29]:

$$\begin{aligned} f(\mathbf{y}|\mathbf{z}, \mathbf{\Theta}) &= h(\mathbf{y}, \mathbf{z})g(\mathbf{\Theta})e^{\delta^\top T(\mathbf{y},\mathbf{z})}, \\ \pi(\mathbf{\Theta}) &= q(\xi, \nu)g(\mathbf{\Theta})^\nu e^{\delta^\top \xi} \end{aligned} \qquad (15)$$

where $\delta \equiv \delta(\mathbf{\Theta})$ is the natural parameter, $T(\mathbf{y}, \mathbf{z})$ is a sufficient statistic of a fixed size, $h(\mathbf{y}, \mathbf{z})$ is a known function, and $g(\mathbf{\Theta})$ is a known normalizing function. Further, ξ is a hyperparameter of the same size as $T(\mathbf{y}, \mathbf{z})$, $\nu \in \mathbb{R}^+$ is a scalar hyperparameter, and $q(\xi, \nu)$ is a known function. Taking advantage of this distribution under conjugacy, the Bayesian update (14) takes the form [30]:

$$\xi_n^h = \xi_n^{h-1} + T(\mathbf{y}_n^h, \mathbf{z}_n^h), \quad \nu_n^h = \nu_n^{h-1} + 1 \qquad (16)$$

For sharing the information in a neighborhood $\mathcal{N}_n$, we use a powerful and robust technique called *diffusion strategy* [31]. One interesting estimation scenario by diffusion strategy run by first exchanging the measurements between the customers during an *adaptation phase*, and after that, exchanging the estimates during a *combination phase* called adapt-then-combine (ATC). The aim of the adaptation phase is to enrich the statistical knowledge of each customer by incorporation of the neighbors' observations. Let us fix other customers' actions and consider customer n turn with its prior density $\pi_n^t(\tilde{\mathbf{\Theta}}^{h-1}|\xi_n^{h-1})$, where ξ_n^{h-1} represents all the information available to customer n by time $h - 1$, which includes its own past observations and those of its neighbors, as well as the parameters of all previous posteriors of the neighbors. Accordingly, the adaptation phase of the diffusion update takes the form:

$$\begin{aligned} \pi_n^t(\tilde{\mathbf{\Theta}}^h|\xi_n^h) &= \pi_n^t(\tilde{\mathbf{\Theta}}^h|\xi_n^{h-1}, \bar{\mathbf{y}}_n^h, \bar{\mathbf{z}}_n^h) \\ &\propto \pi_n^t(\tilde{\mathbf{\Theta}}^{h-1}|\xi_n^{h-1})\prod_{m \in \mathcal{N}_n}[f_m^t(\mathbf{y}_m^h|\mathbf{z}_m^h, \tilde{\mathbf{\Theta}}^{h-1})]^{a_{nm}^h} \end{aligned} \qquad (17)$$

---

[4]A pure strategy profile is said to be a purification of a behavioral/mixed strategy profile if the expected payoffs/distributions of these two strategy profiles are the same for all the players [27].

where $\bar{\boldsymbol{y}}_n^h$ and $\bar{\boldsymbol{z}}_n^h$ are all the new observations and explanatory variables available to customer n at time h, $a_{nm}^h \in \{0, 1\}$ is the adaptation weight assigned by customer n to its neighbor $m \in \mathcal{N}_n$ for adaptation in slot h. If the observation $\boldsymbol{y}_m^h$ is not an outlier, then $a_{nm}^h = 1$; otherwise, $a_{nm}^h = 0$. The hyperparameters of the conjugate prior densities in the adaptation phase of the diffusion protocol are updated as follows:

$$\xi_n^h = \xi_n^{h-1} + \sum_{m \in \mathcal{N}_n} a_{nm}^h T(\boldsymbol{y}_m^h, \boldsymbol{z}_m^h), \; \nu_n^h = \nu_n^{h-1} + \sum_{m \in \mathcal{N}_n} a_{nm}^h \tag{18}$$

consequently, the purpose of the combination phase is to collaboratively improve the estimates of the customers with some proper combination protocols. As we use the KLD function and the customers try to minimize the cumulative loss $\sum_{m \in \mathcal{N}_n} c_{nm}^h D\big(\hat{\pi}_n^t(\tilde{\boldsymbol{\Theta}}^h|\cdot) \| \pi_m^t(\tilde{\boldsymbol{\Theta}}^h|\cdot)\big)$, the combination step to achieve a non-negative minimal KLD takes the form [30]:

$$\hat{\pi}_n^t(\tilde{\boldsymbol{\Theta}}^h|\cdot) \propto \prod_{m \in \mathcal{N}_n} \big[\pi_m^t(\tilde{\boldsymbol{\Theta}}^h|\cdot)\big]^{c_{nm}^h} \tag{19}$$

where coefficients $\{c_{nm,i}, \forall m \in \mathcal{N}_n\}$ are unit $|\mathcal{N}_n|$-simplex weights expressing the degree of belief of customer n in customer $\{m \in \mathcal{N}_n\}$'s information. So, we can formulate the combination phase as:

$$\hat{\xi}_n^h = \sum_{m \in \mathcal{N}_n} c_{nm}^h \xi_m^h, \; \hat{\nu}_n^h = \sum_{m \in \mathcal{N}_n} c_{nm}^h \nu_m^h \tag{20}$$

To compute the combination weights, if we arrange the weights into vector $\boldsymbol{c}_n^h$, it has a categorical distribution parameterized by a probability vector $\boldsymbol{q}_n^h$ with the probability mass function and expected values as [30]:

$$\rho_n^c(\boldsymbol{c}_n^h|\boldsymbol{q}_n^h) = \prod_{m \in \mathcal{N}_n} q_{nm}^h {}^{c_{nm}^h}, \; \mathbb{E}_{\rho_n^c}[c_{nm}^h|\boldsymbol{q}_n^h] = q_{nm}^h \tag{21}$$

where weight c_{nm}^h is determined according to the knowledge about q_{nm}^h. Bayesian estimation of q_{nm}^h with the conjugate Dirichlet prior distribution has the probability density and expectation as follows:

$$\rho_n^q(\boldsymbol{q}_n^h|\boldsymbol{\psi}_n^{h-1}) = \frac{1}{B(\boldsymbol{\psi}_n^{h-1})} \prod_{m \in \mathcal{N}_n} q_{nm}^h {}^{\psi_{nm}^{h-1}-1},$$

$$\mathbb{E}_{\rho_n^q}[q_{nm}^h|\boldsymbol{\psi}_n^{h-1}] = \frac{\psi_{nm}^{h-1}}{\sum_{m \in \mathcal{N}_n} \psi_{nm}^{h-1}} \tag{22}$$

where $B(\cdot)$ is the multivariate beta function and hyperparameters $\{\psi_{nm}^{h-1}, \forall m \in \mathcal{N}_n\}$ are arranged into vector $\boldsymbol{\psi}_n^{h-1}$ to aggregate the knowledge about $\boldsymbol{q}_n^h$. The Bayesian-diffusion update of prior distribution (22) by observations from (21) becomes:

$$\rho_n^q(\boldsymbol{q}_n^h|\boldsymbol{\psi}_n^{h-1}) \propto \prod_{m \in \mathcal{N}_n} q_{nm}^h {}^{c_{nm}^h} \prod_{m \in \mathcal{N}_n} q_{nm}^h {}^{\psi_{nm}^{h-1}-1}$$

$$= \prod_{m \in \mathcal{N}_n} q_{nm}^h {}^{c_{nm}^h + \psi_{nm}^{h-1}-1} \tag{23}$$

However, customer n does not know which neighbor has the best knowledge about $\tilde{\boldsymbol{\Theta}}^t$. So, the customer measures how well neighbor m fits the observations using the predictive performance as follows:

$$\vartheta_{nm}^h = \int \pi_m^t(\tilde{\boldsymbol{\Theta}}^{h-1}|\xi_m^{h-1}) \prod_{j \in \mathcal{N}_n} f_j^t(\boldsymbol{y}_j^h|\boldsymbol{z}_j^h, \tilde{\boldsymbol{\Theta}}^h) d\tilde{\boldsymbol{\Theta}}^h \tag{24}$$

and then, according to the quasi-Bayesian approach uses measure ϑ_{nm}^h in place of c_{nm}^h in (23) [32]. Inspired by the diffusion-Kalman filtering concept [33], one can model the variation of the unknown parameter $\tilde{\boldsymbol{\Theta}}^t$ and the customer n's observation as the following state space model:

$$\tilde{\boldsymbol{\Theta}}^h = A_n^t \tilde{\boldsymbol{\Theta}}^{h-1} + B_n^t \boldsymbol{z}_n^h + \epsilon_n^h,$$

$$\boldsymbol{y}_n^h = H_n^t \tilde{\boldsymbol{\Theta}}^h + \upsilon_n^h \tag{25}$$

where A_n^t, B_n^t, and H_n^t are matrices of compatible dimensions, and both ϵ^h and υ^h are independent and identically distributed Gaussian noises. Let us modify this method under the Bayesian-Kalman filtering approach as follows [34]:

$$\tilde{\boldsymbol{\Theta}}^h|\tilde{\boldsymbol{\Theta}}^{h-1}, \boldsymbol{z}_n^h \sim \mathcal{N}(A_n^t \tilde{\boldsymbol{\Theta}}^{h-1} + B_n^t \boldsymbol{z}_n^h, Q_n^h),$$

$$\boldsymbol{y}_n^h|\tilde{\boldsymbol{\Theta}}^h \sim \mathcal{N}(H_n^t \tilde{\boldsymbol{\Theta}}^h, R_n^h) \tag{26}$$

where Q_n^h and R_n^h are state and observation covariance matrices of customer n. At customer n's side, we denote mean and covariance matrices of the Gaussian prior probability density function $\pi_n^t(\tilde{\boldsymbol{\Theta}}^{h-1}|\boldsymbol{y}_n^{0:h-1}, \boldsymbol{z}_n^{0:h-1})$ by $\bar{\boldsymbol{\Theta}}_n^{-h}$ and $\bar{\Sigma}_n^{-h}$ and the posterior probability density function $\pi_n^t(\tilde{\boldsymbol{\Theta}}^h|\boldsymbol{y}_n^{0:h}, \boldsymbol{z}_n^{0:h})$ by $\bar{\boldsymbol{\Theta}}_n^{+h}$ and $\bar{\Sigma}_n^{+h}$, respectively. At slot h, the customer can predict the value of $\tilde{\boldsymbol{\Theta}}^h$ from $\tilde{\boldsymbol{\Theta}}^{h-1}$ using the state evolution model (26) and Chapman-Kolmogorov equation [34]:

$$\pi_n^t(\tilde{\boldsymbol{\Theta}}^h|\boldsymbol{y}_n^{0:h-1}, \boldsymbol{z}_n^{0:h}) =$$

$$\int \pi_n^t(\tilde{\boldsymbol{\Theta}}^h|\tilde{\boldsymbol{\Theta}}^{h-1}, \boldsymbol{z}_n^h)\pi_n^t(\tilde{\boldsymbol{\Theta}}^{h-1}|\boldsymbol{y}_n^{0:h-1}, \boldsymbol{z}_n^{0:h-1})d\tilde{\boldsymbol{\Theta}}^{h-1}$$

where, the filtering distribution $\pi_n^t(\tilde{\boldsymbol{\Theta}}^{h-1}|\boldsymbol{y}_n^{0:h-1}, \boldsymbol{z}_n^{0:h-1})$ has mean $\bar{\boldsymbol{\Theta}}_n^{+h-1}$ and covariance $\bar{\Sigma}_n^{+h-1}$. The properties of normal distribution ensure that the predicted prior probability density function $\pi_n^t(\tilde{\boldsymbol{\Theta}}^h|\boldsymbol{y}_n^{0:h-1}, \boldsymbol{z}_n^{0:h})$ is also a normal distribution $\mathcal{N}(\bar{\boldsymbol{\Theta}}_n^{-h}, \bar{\Sigma}_n^{-h})$ with hyperparameters:

$$\bar{\boldsymbol{\Theta}}_n^{-h} = A_n^t \bar{\boldsymbol{\Theta}}_n^{+h-1} + B_n^t \boldsymbol{z}_n^h,$$

$$\bar{\Sigma}_n^{-h} = A_n^t \bar{\Sigma}_n^{+h-1}(A_n^t)^\top + Q_n^h \tag{27}$$

Subsequently, Bayes' theorem suggests updating the prior of $\tilde{\boldsymbol{\Theta}}^h$ with the information about $\tilde{\boldsymbol{\Theta}}^h$ in the observed $\boldsymbol{y}_n^h$ and $\boldsymbol{z}_n^h$ as follows:

$$\pi_n^t(\tilde{\boldsymbol{\Theta}}^h|\boldsymbol{y}_n^{0:h}, \boldsymbol{z}_n^{0:h}) = \frac{\pi_n^t(\tilde{\boldsymbol{\Theta}}^h|\boldsymbol{y}_n^{0:h-1}, \boldsymbol{z}_n^{0:h})f_n^t(\boldsymbol{y}_n^h|\tilde{\boldsymbol{\Theta}}^h)}{\int \pi_n^t(\tilde{\boldsymbol{\Theta}}^h|\boldsymbol{y}_n^{0:h-1}, \boldsymbol{z}_n^{0:h})f_n^t(\boldsymbol{y}_n^h|\tilde{\boldsymbol{\Theta}}^h)d\tilde{\boldsymbol{\Theta}}^h} \tag{28}$$

where, from (26) we can write the observation model as:

$$f(\boldsymbol{y}_n^h|\tilde{\boldsymbol{\Theta}}^h) \propto$$

$$exp\{-\frac{1}{2}(\boldsymbol{y}_n^h - H_n^t \tilde{\boldsymbol{\Theta}}^h)^\top (R_n^h)^{-1}(\boldsymbol{y}_n^h - H_n^t \tilde{\boldsymbol{\Theta}}^h)\} =$$

$$exp\Big\{\mathrm{Tr}\Big(-\frac{1}{2}\underbrace{\begin{bmatrix} -1 \\ \tilde{\boldsymbol{\Theta}}^h \end{bmatrix}\begin{bmatrix} -1 \\ \tilde{\boldsymbol{\Theta}}^h \end{bmatrix}^\top}_{\delta}\underbrace{\begin{bmatrix} (\boldsymbol{y}_n^h)^\top \\ (H_n^t)^\top \end{bmatrix}(R_n^h)^{-1}\begin{bmatrix} (\boldsymbol{y}_n^h)^\top \\ (H_n^t)^\top \end{bmatrix}^\top}_{T(\boldsymbol{y}_n^h)}\Big)\Big\}$$

with compilable form of conjugate normal distribution as:

$$\pi_n^t(\tilde{\boldsymbol{\Theta}}^h|\boldsymbol{y}_n^{0:h-1}, \boldsymbol{z}_n^{0:h}) \propto$$

$$exp\{-\frac{1}{2}(\tilde{\boldsymbol{\Theta}}^h - \bar{\boldsymbol{\Theta}}_n^{-h})^\top (\bar{\Sigma}_n^{-h})^{-1}(\tilde{\boldsymbol{\Theta}}^h - \bar{\boldsymbol{\Theta}}_n^{-h})\} =$$

$$exp\Big\{\mathrm{Tr}\Big(-\frac{1}{2}\underbrace{\begin{bmatrix} -1 \\ \tilde{\boldsymbol{\Theta}}^h \end{bmatrix}\begin{bmatrix} -1 \\ \tilde{\boldsymbol{\Theta}}^h \end{bmatrix}^\top}_{\delta} \underbrace{\begin{bmatrix} (\bar{\boldsymbol{\Theta}}_n^{-h})^\top \\ I \end{bmatrix} (\bar{\Sigma}_n^{-h})^{-1}\begin{bmatrix} (\bar{\boldsymbol{\Theta}}_n^{-h})^\top \\ I \end{bmatrix}^\top}_{\xi_n^h}\Big)\Big\}$$

So, we can replace Bayesian update (28) by updating the hyperparameters (16) as:

$$\xi_n^h = \xi_n^{h-1} + T(\boldsymbol{y}_n^h, \boldsymbol{z}_n^h) = \begin{bmatrix} \xi_n^{11} & \xi_n^{12} \\ \xi_n^{21} & \xi_n^{22} \end{bmatrix} \tag{29}$$

$$\xi_n^{11} = (\bar{\boldsymbol{\Theta}}_n^{-h})^\top (\bar{\Sigma}_n^{-h})^{-1}\bar{\boldsymbol{\Theta}}_n^{-h} + (\boldsymbol{y}_n^h)^\top (R_n^h)^{-1}\boldsymbol{y}_n^h, \tag{30}$$

$$\xi_n^{12} = (\bar{\boldsymbol{\Theta}}_n^{-h})^\top (\bar{\Sigma}_n^{-h})^{-1} + (\boldsymbol{y}_n^h)^\top (R_n^h)^{-1}H_n^t, \tag{31}$$

$$\xi_n^{21} = (\bar{\Sigma}_n^{-h})^{-1}(\bar{\boldsymbol{\Theta}}_n^{-h})^\top + (H_n^t)^\top (R_n^h)^{-1}\boldsymbol{y}_n^h, \tag{32}$$

$$\xi_n^{22} = (\bar{\Sigma}_n^{-h})^{-1} + (H_n^t)^\top (R_n^h)^{-1}H_n^t \tag{33}$$

Now, applying the adaptation phase (18) leads to [34]:

$$\xi_n^h = \xi_n^{h-1} + \sum_{m\in\mathcal{N}_n} a_{nm}^h \begin{bmatrix} \boldsymbol{y}_m^h \\ H_m^t \end{bmatrix}(R_m^h)^{-1}\begin{bmatrix} \boldsymbol{y}_m^h \\ H_m^t \end{bmatrix}^\top \tag{34}$$

$$\nu_n^h = \nu_n^{h-1} + |\mathcal{N}_n| \tag{35}$$

and the classical Kalman filter update equations are drown using the least-squares estimator, exploiting the blocks of the matrix ξ_n^h in (29) and applying the adaptation phase (similar to (34)) on this matrix [30], [34]:

$$\bar{\Sigma}_n^{+h} = \sum_{m\in\mathcal{N}_n} a_{nm}^h [\xi_n^{22}]^{-1} =$$

$$\Big[(\bar{\Sigma}_n^{-h})^{-1} + \Big(\sum_{m\in\mathcal{N}_n} a_{nm}^h (H_m^t)^\top (R_m^h)^{-1}H_m^t\Big)\Big]^{-1},$$

$$\bar{\boldsymbol{\Theta}}_n^{+h} = \sum_{m\in\mathcal{N}_n} a_{nm}^h [\xi_n^{22}]^{-1}\xi_n^{21} =$$

$$\sum_{m\in\mathcal{N}_n} a_{nm}^h \bar{\Sigma}_n^{+h}\Big[(\bar{\Sigma}_n^{-h})^{-1}(\bar{\boldsymbol{\Theta}}_n^{-h})^\top + (H_m^t)^\top (R_m^h)^{-1}\boldsymbol{y}_m^h\Big] =$$

$$\bar{\boldsymbol{\Theta}}_n^{-h} + \bar{\Sigma}_n^{+h}\Big[\sum_{m\in\mathcal{N}_n} a_{nm}^h (H_m^t)^\top (R_m^h)^{-1}(\boldsymbol{y}_m^h - H_m^t\bar{\boldsymbol{\Theta}}_n^{-h})\Big]$$

Applying the combination phase (20) to the above equation results in the desired parameters:

$$\hat{\Sigma}_n^h = \Big[\sum_{m\in\mathcal{N}_n} c_{nm}^h (\tilde{\Sigma}_m^{+h})^{-1}\Big]^{-1}$$

$$\hat{\boldsymbol{\Theta}}_n^h = \hat{\Sigma}_n^h\Big(\sum_{m\in\mathcal{N}_n} c_{nm}^h (\bar{\Sigma}_m^{+h})^{-1}\bar{\boldsymbol{\Theta}}_m^{+h}\Big) \tag{36}$$

where mean $\hat{\boldsymbol{\Theta}}_n^h$ and covariance $\hat{\Sigma}_n^h$ are used to estimate $\tilde{\boldsymbol{\Theta}}^h$. Finally, we present our Bayesian-diffusion-Kalman filtering ECS mechanism (BdiffKF-ECS) in Algorithm 1 which is executed by each PEV n.

Theorem 2. *The proposed mechanism in Algorithm 1 converges to a unique PSNE with a finite number of iterations if the power grid topology satisfies the following conditions:*

Condition **P**: *For each player $n \in \mathcal{N}$, each neighbor $j \in \mathcal{N}_n$, any strategies $\boldsymbol{X}_j^h, \acute{\boldsymbol{X}}_j^h \in \mathcal{X}_j^h$, $\boldsymbol{X}_n^h \in \mathcal{X}_n^h$ and*

Algorithm 1 BdiffKF-ECS Mechanism

1: I. **Initialization:** Set X_n^0, $[\vartheta_{nm}^0]_{m\in\mathcal{N}_n}$, $\hat{\boldsymbol{\Theta}}_n^0$, the convergence criteria ϵ, and $\hat{\Sigma}_n^0$.
2: II. **Repeat** for $t = 1, 2, \cdots$:
3:　　Take parameters of previous period $t - 1$ as initial values and get H_m^t of neighbors $m \in \mathcal{N}_n$.
4:　　**Repeat** for $h = 1, \cdots, H^t$:
5:　　　Let $\bar{\boldsymbol{\Theta}}_n^{-h} = \hat{\boldsymbol{\Theta}}_n^{h-1}$ and $\bar{\Sigma}_n^{-h} = \hat{\Sigma}_n^{h-1}$.
6:　　　**Estimation Phase:**
7:　　　　Prediction Step: Run iterations (27).
8:　　　　Adaptation Step: Get $\boldsymbol{y}_m^h$ and R_m^h of neighbors and run iterations (36).
9:　　　　Combination Step: Get $\bar{\boldsymbol{\Theta}}_m^{+h}$ and $\bar{\Sigma}_m^{+h}$ of neighbors, update the Dirichlet prior hyperparameters ψ_n^h of (23) and determine $[c_{nm}^h]_{m\in\mathcal{N}_n}$ from (24), and run iterations (36)
10:　　　**Action Phase:**
11:　　　　**Repeat** for $i = 1, 2, \cdots$ according to some order:
12:　　　　　Receive last update $\boldsymbol{X}_{m,i}^h$ and $\boldsymbol{l}_{m,i}^h$ of neighbors.
13:　　　　　Solve (11) and determine $\boldsymbol{X}_{n,i+1}^h$.
14:　　　　　**If** $|\boldsymbol{X}_{n,i+1}^h - \boldsymbol{X}_{n,i}^h| \geq \epsilon$ (i.e., the new schedule changes compared to the current schedule), **Then** set $\boldsymbol{X}_{n,i+1}^h$ as the new solution and broadcast it to the neighbors.
15:　　　　**Until convergence (i.e., none of the customers broadcast their schedules)**
16:　　　　Customer n applies element $\boldsymbol{x}_n^h$ of $\boldsymbol{X}_n^h$ and discards the others according to the rolling horizon (e.g., [37]).

$\boldsymbol{X}_{\mathcal{N}_n/j}^h \in \mathcal{X}_{\mathcal{N}_n/j}^h$, *we must have* $U_n(\boldsymbol{X}_n^h, \boldsymbol{X}_j^h, \boldsymbol{X}_{\mathcal{N}_n/j}^h) \neq U_n(\boldsymbol{X}_n^h, \acute{\boldsymbol{X}}_j^h, \boldsymbol{X}_{\mathcal{N}_n/j}^h)$.

Condition **N**: *For each player $n \in \mathcal{N}$, for any neighbors $j, i \in \mathcal{N}_n$ 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 n which passes neither through j nor i.*

Proof. See Appendix B.　　□

Condition P implies that the action of a players always affects his neighbors' payoffs, so, its behavior is learnable (deviation is detectable). A 3-connected[5] 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 [35] for detailed information and proof about Conditions **P** and **N**). To guarantee the power grid topology satisfies Condition **N**, one can let the customers decide their neighbors. Then the utility company check if the graph is connected (has the Laplacian matrix with positive second smallest eigenvalue [36]) and each customer has at least 3 customers (the degree of the graph is at least 3). Otherwise, with some manipulation, the utility company can fix the system topology.

V. NUMERICAL RESULTS

A. Simulation Setup

For the simulation scenario, it is assumed that there is one utility company procuring electricity from both conventional and renewable resources and serving 50-2000 customers of different types (e.g., residential, commercial, educational, organizational, and industrial). Each day $t \in \mathcal{T}$ is divided into H^t = 96 slots each with 15 min duration. The baseline day-ahead

---

[5]Graph 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 (and the corresponding links) have been removed. Simply, a graph is n-connected if any two of its nodes can be joined by n independent paths.

electricity price data is drawn randomly from Pennsylvania-New Jersey-Maryland Interconnection (PJM) electricity market [38]. It is assumed that each customer has one PEV with charge/discharge power rate $2 \leq x_n^{rat} \leq 4$, efficiency $\eta_{n,c}, \eta_{n,d} = 1$, and desire energy demand randomly within $E_n^{lb} = 6$ and $E_n^{ub} = 12$. The battey capacity of the PEVs are chosen randomly between [15-20] with uniform distribution. The customers' base loads for each hour are adopted randomly between 3 and 10 kW with some correlation with the baseline price signal. For each customer additive noise power ϵ_n^h and υ_n^h are randomly added between (23-69) dB and (69-92) dB, respectively. The aggregate power output of the renewable resources is randomly selected with Gaussian distribution $\tilde{P}_{k,r}^h \sim \mathcal{N}(\bar{P}_{k,r}^h, \sigma_{k,r}^2)$ and the regulation signal is considered as $\tilde{w}^h = \tilde{P}_{k,r}^h - \bar{P}_{k,r}^h$.

The system topology (graph) for the simulations are generated by putting randomly customers in the square region $[0,4] \times [0,4]$ km^2 in a way to satisfies the conditions in Theorem 2. Customers that are within a certain radius (e.g., within radius = 0.5 km) from each other are declared neighbors and connected to each other. To guarantee this, we repeat generating random topologies with increasing search radius until the degree of each vertex in the acquired topology is no less than 3 and the second smallest eigenvalue of the Laplacian matrix is positive. The combination weights c_{nm}^h are initialized according to the Metropolis rule introduced in [39], and then replaced by weights ϑ_{nm}^h updated according to (24) at each slot h. The simulations are done in PC Laptop 64-bit Intel(R) Core(TM) i7-4510U CPU @ 2.00-2.60GHz RAM 8.00GB with MATLAB R2017b and its CVX toolbox working with Gurobi.

B. Performance Results

The performance of the estimation part in Algorithm 1 is shown in Fig. 2. As a measure of estimation performance between Non-Cooperative mode (in which there is not any information sharing) and Diffusion mode (in which the customers cooperate and share information with each other), Fig. 2(a) denotes the average mean square deviation (MSD) from the true value. Clearly, in the non-cooperative case the estimation performance is much worse and more unreliable. The learning and tracking abilities of the proposed method are demonstrated in Figs. 2(b) and (c) for estimating the baseline and regulation signals, respectively. As an insight, we can see that after a few iterations the estimated values match the true value with acceptable accuracy. The convergence of the proposed method in solving problem (10) is analyzed in Fig. 3. As shown by the results of Figs. 3(a) and (b), only one iteration per customer (total of 50 iterations per 50 customers) suffices for convergence. So, we can consider the proposed mechanism as a one-shot infinitely repeated Bayesian game. Further, both figures show that the customers' payment depends on the parameters γ^h in (6). So, the utility company can change the customers' consumption pattern as his wish by properly tuning γ^h.

For better presenting the impact of γ^h on the customers' behavior, Figs. 4(a) and (b) are provided. In Fig. 4(a), the
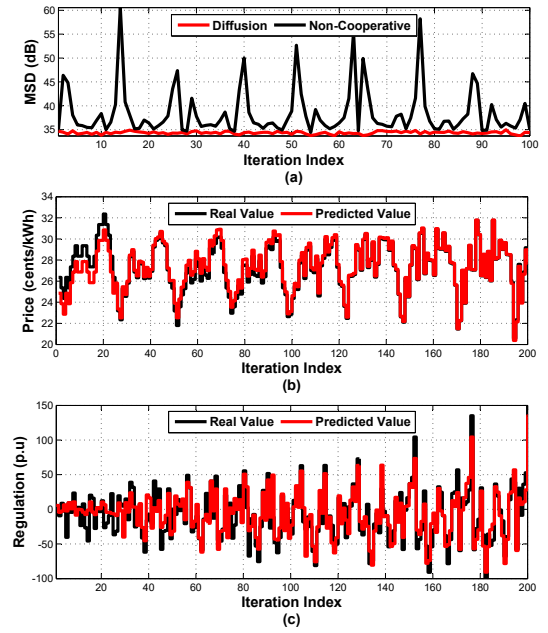

Fig. 2. Analysis of the estimation part; (a) total average mean square deviation (MSD), (b) price estimation and tracking performance, (c) estimation and tracking of the regulation signal.
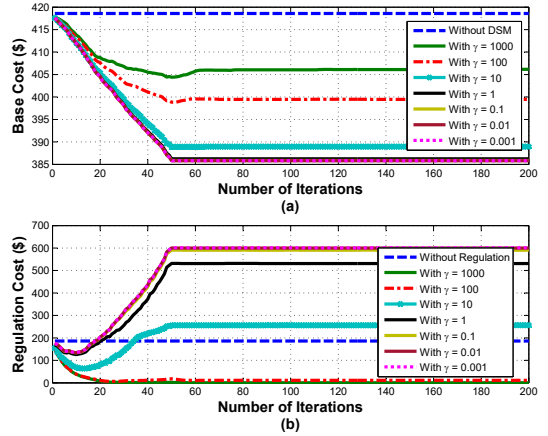

Fig. 3. Value and convergence comparison of the total system payment with tuning γ^h; (a) in terms of the total baseline price $\tilde{p}_b^h$, and (b) total shadow price $\tilde{p}_{sh}^h$.

parameters of the system are demonstrated. The behavior of aggregated charge/discharge of the PEVs is denoted in Fig. 4(b). As depicted, when the shadow price is much higher than the baseline price, the customers attempt to track the regulation signal as much as possible to reduce the total payment. However, while γ^h reduces, the impact of the baseline price on the total payment is increased. So, the customers try to charge their PEVs in slots with low baseline prices and discharge them at slots with high prices.

The applications of the proposed framework in terms of the two important aspects of power systems, namely, ancillary service and load curve shaping (e.g., valley filling and pick clipping), are analyzed and the results are denoted in Figs. 5(a) and (b). Fig. 5(a) shows the ancillary service result in which the total consumption of the PEVs tracks the regulation signal

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2866267, IEEE Transactions on Industrial Informatics
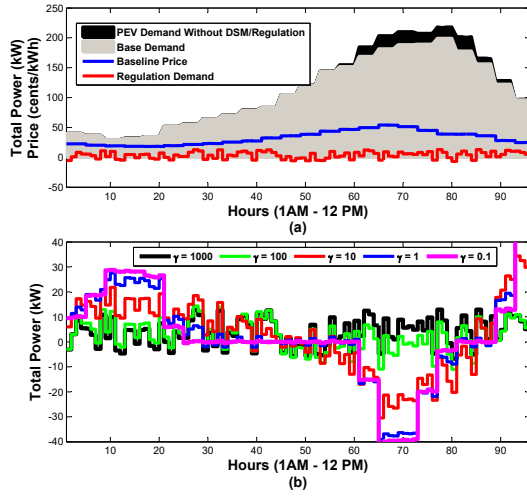
10

Fig. 4. Changing the total system behavior by manipulating the real-time price; (a) the system data, and (b) the process of changing the charge/discharge of the PEVs.
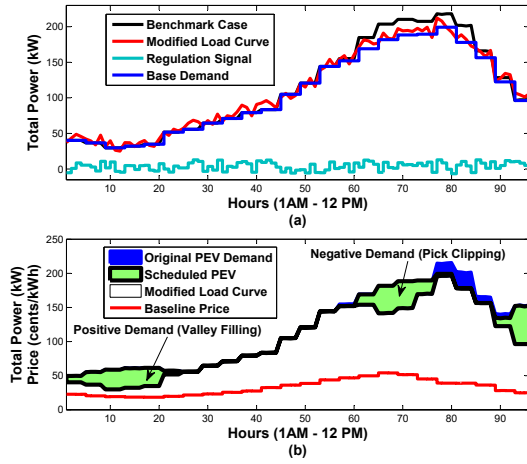

Fig. 5. The applications of the proposed BdiffKF-ECS mechanism; (a) ancillary service, and (b) shaping the curve of the total load demand.

to mitigate the destructive effects of power fluctuations of the renewable resources. The result of the proposed mechanism in Algorithm. 1 under the DSM protocol is depicted in Fig. 5(b). We can see that at slots with low demand the PEVs are in charge mode (to fill the valley) and at the slots with high (peak)-demand are in discharge mode (to curtail the peak). However, as shown, the discharge time of the PEVs are not effectively coincided with the true peak, because the baseline price is not appropriately determined by the utility company. This shows the importance of adjusting the baseline price signal to having the highest value in the highest demand time.

As we formulated our problem as an infinitely repeated Bayesian game, it is necessary to evaluate its the long-term performance. The signal tracking ability of Algorithm 1 is denoted in Fig. 6. Clearly, the performance of the algorithm in tracking the frequency regulation signal coming from the utility company is acceptable. We must note that none of the customer's constraints in (1) is violated in this result and all the PEVs have reached to their desirable charge level
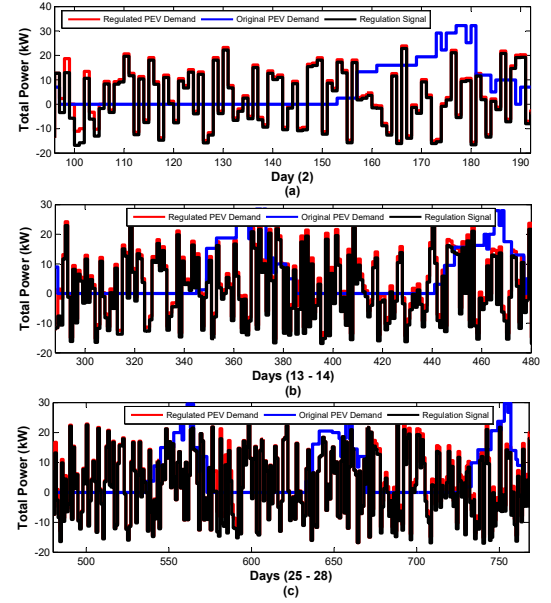

Fig. 6. The long-terms performance of the proposed mechanism over 30 days; this is in terms of the frequency regulation application.
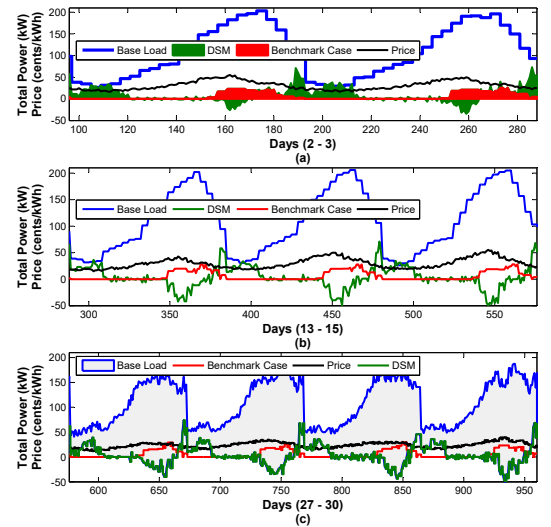

Fig. 7. The long-terms performance of the proposed mechanism over 30 days; this is in terms of the DSM application.

$E_n^{d,t}$ at the end of each day. In another scenario, the long-term performance of the DSM application is analyzed over 30 days. The results are denoted in Fig. 7, which show that the proposed repeated game works well and fills/curtails the aggregate demand valleys/peaks.

In modeling the interactions between the customers, the game theoretic methods turns out to be the state of the art dominant strategy in the DSM and PEV charge/discharge scheduling literature [7], [8], [13], [40], [41]. To guarantee the convergence of these methods to an equilibrium (e.g., Nash equilibrium), it is essential for each customers to know the consumption pattern of all other customers in the system (either directly or through a third party entity). Further, to achieve a global optimal solution, it is necessary for the system

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2018.2866267, IEEE Transactions on Industrial Informatics

11

to determine a proper price signal and broadcast it to the customers step-by-step. In our real-time scheduling program, with growing the system, this procedure would impose high communication and computational burden to the system as denoted in Fig. 8(a). From this results we can find out that for a wide system with 2000 customer, it took about 41 min for the game theoretic mechanism with a centralized price determining mechanism (Nash Game (Centralized)) to converge to an optimal solution, which is not acceptable. Estimating the price signal with our diffusion framework would reduced the computation/communication time of the game theoretic mechanism (Nash Game (Diffusion)) to almost 33, which is still long time for the real-time application. However, Fig. 8(a) shows that our BdiffKF-ECS mechanism has very low computation/communication time, specially for the BdiffKF-ECS (Diffusion) scenario which takes 8 min.

The most important thing is that even with such a low computation/communication between the customer, the optimal solution of the proposed framework is almost the same as optimal solution of the Nash games. The results of Fig. 8(b) asserts our claim. From this figure we can see that the centralized solution achieves the minimum aggregate cost. However, the centralized solutions imposes a huge computational burden to such a with system and is not applicable in real world. Beside, as in centralized solutions all the data are gathered in the central controller, the privacy of the customers is always in danger. Another drawback of the centralized solutions is that any lack of proper operation in the central control (either deliberately because of the cyber-attacks or unintentionally because of communication failures and inaccuracies) have high destructive effect on the system efficiency. In the stochastic problems with too many stochastic parameters like ours, these drawbacks put down the viability of the solution mechanism as denoted in Fig. 9(a). As we can see in this figure, inaccuracy (noise) in determining the stochastic parameters $\tilde{p}^h$ and $E_n^{d,t}$ as well as random link failure between some of the customers make the "Nash Game (Centralized)" mechanism unstable and non-optimal. However, when the price signal $\tilde{p}^h$ is estimated through our Bayesian-diffusion framework "Nash Game (Diffusion)", the mechanism becomes more stable and the aggregate cost reduces. The most effective case is when both the scheduling part and price estimation part are implemented according to our mechanism "BdiffKF-ECS (Diffusion)". In this case, the proposed mechanism is robust to link failure and noise (as the communications between the customers are done locally) and works well in the presence of the uncertainty resources (tanks to our real-time adaptive stochastic mechanism).

The DSM performance comparison results of the proposed framework under uncertainty is demonstrated in Fig. 9(b). Considering the uncertainty nature of customers deriving pattern (i.e., parameters α_n^t, β_n^t, and $E_n^{d,t}$) results in lower peak demand (case "Load Curve(Our Mechanism)") and better utilizing the potential flexibility of the PEVs (case "PEV(Our Mechanism)"). In the case "PEV(No Learning)" in which the classical game is implemented instead of the Bayesian repeated game, the high charging(discharging) power is not properly coincided with the low(high) price value. This can
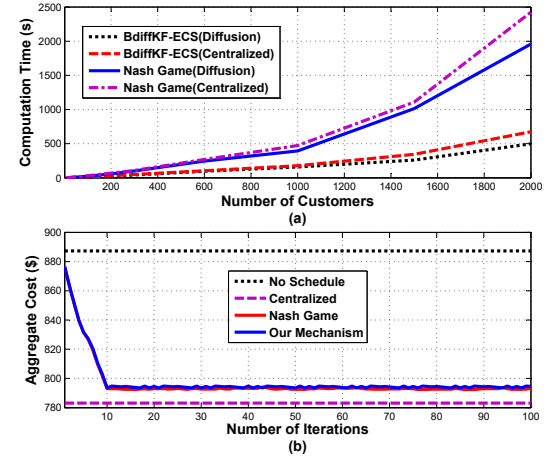

Fig. 8. Performance comparison between the proposed and the Nash game approaches; (a) computation time, and (b) total payment.
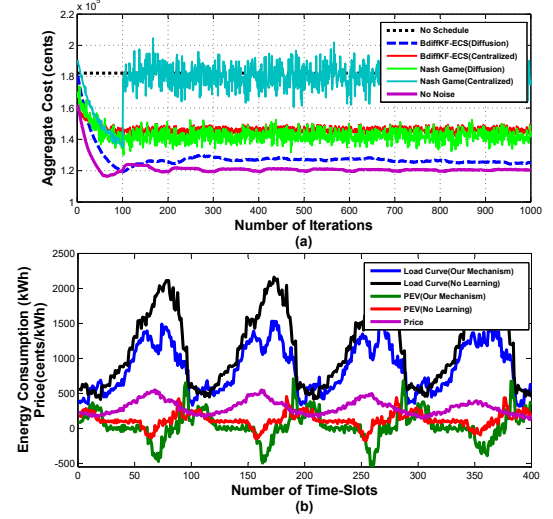

Fig. 9. Performance comparison between the proposed and the Nash game approaches in the presence of the uncertainty sources; (a) Aggregate cost, and (b) Consumption pattern.

increase the customers payment and the total system peak demand compared to our mechanism.

VI. CONCLUSIONS

In this paper, we proposed a universal real-time distributed DSM-ancillary service framework applied to the PEVs using the Bayesian repeated game, the diffusion strategy, and the Kalman filtering theory. We characterized the PEVs specification and provided a pure strategic game using the rolling horizon theory to achieve an optimal solution and to fully take the customers' attribution into account. Further, we formulated a novel real-time pricing policy by which the utility company can control and modify the consumption pattern of the PEVs to mitigate the adverse effect of fluctuation of the RERs and charge of the PEVs. The simulation results demonstrated that by appropriately determining the price and regulation signals, the utility company can increase the reliability of the power

system by making the balance between the generated and consumed power.

APPENDIX

A. Proof of Theorem 1

$\star_1$: Let $\mathscr{C}_n^h$ be a sub-σ-algebra of $\mathscr{S}_n^h$. Clearly, for each $h \in \mathcal{H}^t$, the local objective function U_n^t in (11) is $\mathscr{B}(\mathcal{X}^h) \otimes \mathscr{C}_n^h$-measurable on $\mathcal{X}^h \times \mathcal{S}_n^h$, is continuous on $\mathcal{X}^h$ for all feasible charging profiles $\boldsymbol{X}_n \in \mathcal{X}_n^h$, and is integrably bounded[6].

$\star_2$: The private information $\mathcal{S}_n^h$ of each player $n \in \mathcal{N}$ in the proposed game model is independent[7] of all other players' private information.

$\star_3$: For each customer $n \in \mathcal{N}$, $\mathscr{S}_n^h$ is atomless[8] and $\mathscr{C}_n^h$ is setwise coarser[9] than $\mathscr{S}_n^h$.

Proposition 2. *From $\star_1$, $\star_2$, and $\star_3$ we can claim that every $\mathscr{C}^h = \otimes_{n=1}^N \mathscr{C}_n^h$-measurable behavioral strategy profile of our Bayesian game has a $\mathscr{S}^h = \otimes_{n=1}^N \mathscr{S}_n^h$-measurable undistinguishable purification (see Theorem 2 of [42] for proof).*

Proposition 3. *Features $\star_1$, $\star_2$, and $\star_3$ provide the necessary and sufficient conditions (i.e., they guarantee a saturated probability space for each customer[10]) for the existence of a PSNE in our model (see Theorem 1 of [42] for proof).*

Using the backward induction, the rolling horizon theory implies that the problem at hand at each slot h has an equivalent pure strategy and sub-sequentially a PSNE [28].

B. Proof of Theorem 2

The proposed mechanism in Algorithm 1 is comprised of infinite sub-games repeated at each slot $h \in \mathcal{H}$ of each period $t \in \mathcal{T}$. The convergence analysis must be done using backward induction denoting that each sub-game in $h \in \mathcal{H}^t \subset \mathcal{T}$ results in a subgame perfect equilibrium. However, we analyze one sub-game and the extension to the infinitely repeated game is straightforward using backward induction [44]. The algorithm convergence comprises of three aspects:

$\star_1$ **Convergence of the Bayesian-diffusion estimation procedure:** In [45], it has been proved that by adopting appropriate adaptation and combination weights, the diffusion estimation process converges with good accuracy and the benefit of this cooperation is examined in detail in [46]. Similarly, the convergence of the Bayesian version of diffusion strategy has been proved in [30]. By adopting proper outlier

analysis for choosing a_{nm}^h according to work in [47], and proper approach for choosing a_{nm}^h, such as the quasi-Bayesian [48] or adaptive combination weights [49], the convergence of the estimation stage is guaranteed.

$\star_2$ **Convergence of the optimization procedure:** At each iteration i customer n solves the optimization problem (11). We solved this problem using mixed-integer disciplined convex programming methods, such as branch and bound and sequential quadratic programming, for which the convergence to a unique optimal/sub-optimal solution has been proved [50].

$\star_3$ **Convergence of the game-theoretic framework:** The payoff function in (11) is strictly monotone with respect to its arguments and satisfies condition **P**. Considering an appropriate topology satisfying condition **N** each sub-game at slot τ in scheduling window $\mathcal{H}_n^\tau \subseteq \mathcal{H}$ converges to a unique Nash-Folk equilibrium according to Theorem 2.12 of [35], which is the subgame perfect equilibrium of our Bayesian repeated game.

The objective function of each customer n proposed in (11) is strictly concave. So, the proposed sub-game at each slot becomes a strictly concave N-person game with unique subgame perfect equilibrium according to Theorem 2 of [51]. This unique Nash equilibrium of considered game is the optimal solution of the global minimization problem (10) according to Theorem 2 of [40]. Finally, using $\star_1$, $\star_2$, $\star_3$, backward induction theory, and the results of Theorem 1 we can conclude that our game converges to a unique PSNE.

---

[6]If there is a real-valued integrable function ϕ_n on $\mathcal{S}_n^h, \mathscr{C}_n^h, \pi_n^t$, such that $|U_n^t(a, x_n)| \leq \phi_n^t(x_n)$ for all $(a, x_n) \in \mathcal{X}^h \times \mathcal{S}_n^h$, then U_n^t is integrably bounded

[7]This assumption can be weakened and correlations are allowed.

[8]Given measurable space $(\mathcal{S}_n^h, \mathscr{S}_n^h)$ and a measure μ on this space, a set $\mathcal{A} \subset \mathcal{S}_n^h$ in $\mathscr{S}_n^h$ is called an atom if $\mu(\mathcal{A}) > 0$ and for any measurable subset $\mathcal{B} \subset \mathcal{A}$ with $\mu(\mathcal{B}) < \mu(\mathcal{A})$, set $\mathcal{B}$ has measure zero. A space which has no atom is called atomless (non-atomic) [15].

[9]$\mathscr{C}_n^h$ is setwise coarser than $\mathscr{S}_n^h$ if for every $\mathcal{A} \in \mathscr{S}_n^h$ with positive measure π_n, there exists a $\mathscr{S}_n^h$-measurable subset $\mathcal{A}_1 \subseteq \mathcal{A}$ such that $\pi_n^t(\mathcal{A}_1 \triangle \mathcal{A}_2) > 0$ for any $\mathcal{A}_2 \in \mathscr{C}_n^{\mathcal{A}}$, where $\mathscr{C}_n^{\mathcal{A}}$ is the σ-algebra $\{\mathcal{A} \bigcap \bar{\mathcal{A}} : \bar{\mathcal{A}} \in \mathscr{C}_n^h\}$.

[10]In [43] it has been proved that for non-saturated (but atomless) probability spaces, there may not be any PSNE.

REFERENCES

[1] "Ev volumes.com," *The electric vehicle world sales database*, 2017. [Online]. Available: http://www.ev-volumes.com/

[2] W. Hu, C. Su, Z. Chen, and B. Bak-Jensen, "Optimal operation of plug-in electric vehicles in power systems with high wind power penetrations," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 3, pp. 577–585, 2013.

[3] C. Jin, X. Sheng, and P. Ghosh, "Optimized electric vehicle charging with intermittent renewable energy sources," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 6, pp. 1063–1072, 2014.

[4] T. Zhang, W. Chen, Z. Han, and Z. Cao, "Charging scheduling of electric vehicles with local renewable energy under uncertain electric vehicle arrival and grid power price," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 6, pp. 2600–2612, 2014.

[5] C. Chen and S. Duan, "Optimal integration of plug-in hybrid electric vehicles in microgrids," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 3, pp. 1917–1926, 2014.

[6] 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, 2015.

[7] N. Y. Soltani, S.-J. Kim, and G. B. Giannakis, "Real-time load elasticity tracking and pricing for electric vehicle charging," *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1303–1313, 2015.

[8] W. Lee, L. Xiang, R. Schober, and V. W. Wong, "Electric vehicle charging stations with renewable power generators: A game theoretical analysis," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 608–617, 2015.

[9] P. Kou, D. Liang, L. Gao, and F. Gao, "Stochastic coordination of plug-in electric vehicles and wind turbines in microgrid: A model predictive control approach," *IEEE Transactions on Smart Grid*, vol. 7, no. 3, pp. 1537–1551, 2016.

[10] R. Wang, P. Wang, and G. Xiao, "Two-stage mechanism for massive electric vehicle charging involving renewable energy," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 6, pp. 4159–4171, 2016.

[11] Y. Yang, Q.-S. Jia, G. Deconinck, X. Guan, Z. Qiu, and Z. Hu, "Distributed coordination of ev charging with renewable energy in a microgrid of buildings," *IEEE Transactions on Smart Grid*, 2017.

[12] H. Yang, H. Pan, F. Luo, J. Qiu, Y. Deng, M. Lai, and Z. Y. Dong, "Operational planning of electric vehicles for balancing wind power and load fluctuations in a microgrid," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 2, pp. 592–604, 2017.

[13] C. Luo, Y.-F. Huang, and V. Gupta, "Stochastic dynamic pricing for ev charging stations with renewables integration and energy storage," *IEEE Transactions on Smart Grid*, 2017.

[14] L. P. Hansen and T. J. Sargent, "Robust control and model uncertainty," *The American Economic Review*, vol. 91, no. 2, pp. 60–66, 2001.

[15] H. Federer, *Geometric measure theory*. Springer, 2014.

[16] T. Ackermann, *Wind power in power systems*. John Wiley & Sons, 2005.

[17] F. Luo, J. Zhao, J. Qiu, J. Foster, Y. Peng, and Z. Dong, "Assessing the transmission expansion cost with distributed generation: An australian case study," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1892–1904, 2014.

[18] M. R. Patel, *Wind and solar power systems: design, analysis, and operation*. CRC press, 2005.

[19] D. R. Cox, *The theory of stochastic processes*. Routledge, 2017.

[20] J. L. Doob and J. L. Doob, *Stochastic processes*. Wiley New York, 1953, vol. 7, no. 2.

[21] S. Assoudou and B. Essebbar, "A bayesian model for binary markov chains," *International Journal of Mathematics and Mathematical Sciences*, vol. 2004, no. 8, pp. 421–429, 2004.

[22] W. R. Gilks, S. Richardson, and D. Spiegelhalter, *Markov chain Monte Carlo in practice*. CRC press, 1995.

[23] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 52–63, 2013.

[24] M. J. Osborne and A. Rubinstein, *A course in game theory*. MIT press, 1994.

[25] G. Carmona, *Existence and stability of Nash equilibrium*. World Scientific, 2013.

[26] O. Carbonell-Nicolau and R. P. McLean, "On the existence of nash equilibrium in bayesian games," *Mathematics of Operations Research*, 2017.

[27] P. Prokopovych and N. C. Yannelis, "On the existence of mixed strategy nash equilibria," *Journal of mathematical economics*, vol. 52, pp. 87–97, 2014.

[28] D. Bredstrm, P. Flisberg, and M. Rnnqvist, "A new method for robustness in rolling horizon planning," *International Journal of Production Economics*, vol. 143, no. 1, pp. 41 – 52, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S092552731100065X

[29] C. M. Bishop, *Pattern recognition and machine learning*. springer, 2006.

[30] K. Dedecius and P. M. Djurić, "Sequential estimation and diffusion of information over networks: A bayesian approach with exponential family of distributions," *IEEE Transactions on Signal Processing*, vol. 65, no. 7, pp. 1795–1809, 2017.

[31] F. S. Cattivelli, C. G. Lopes, and A. H. Sayed, "Diffusion recursive least-squares for distributed estimation over adaptive networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 1865–1877, 2008.

[32] M. Kárnỳ, J. Kadlec, and E. L. Sutanto, "Quasi-bayes estimation applied to normal mixture," in *Preprints of the 3rd European IEEE Workshop on Computer-Intensive Methods in Control and Data Processing*. Praha, 1998, pp. 77–82.

[33] F. S. Cattivelli and A. H. Sayed, "Diffusion strategies for distributed kalman filtering and smoothing," *IEEE Transactions on automatic control*, vol. 55, no. 9, pp. 2069–2084, 2010.

[34] K. Dedecius, "Diffusion estimation of state-space models: Bayesian formulation," in *Machine Learning for Signal Processing (MLSP), 2014 IEEE International Workshop on*. IEEE, 2014, pp. 1–6.

[35] M. Laclau, "A folk theorem for repeated games played on a network," *Games and Economic Behavior*, vol. 76, no. 2, pp. 711 – 737, 2012.

[36] R. Merris, "Laplacian matrices of graphs: a survey," *Linear algebra and its applications*, vol. 197, pp. 143–176, 1994.

[37] R. Palma-Behnke, C. Benavides, F. Lanas, B. Severino, L. Reyes, J. Llanos, and D. Sáez, "A microgrid energy management system based on the rolling horizon strategy," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 996–1006, 2013.

[38] *Pennsylvania-New Jersey-Maryland Interconnection (PJM) electricity market*, 2017. [Online]. Available: http://www.pjm.com/markets-and-operations/energy/day-ahead/lmpda.aspx

[39] M. Latifi, A. Khalili, A. Rastegarnia, and S. Sanei, "Fully distributed demand response using the adaptive diffusion–stackelberg algorithm,"

[40] 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.

[41] W. Saad, Z. Han, H. V. Poor, and T. Basar, "Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 86–105, 2012.

[42] W. He and X. Sun, "On the diffuseness of incomplete information game," *Journal of Mathematical Economics*, vol. 54, pp. 131–137, 2014.

[43] M. A. Khan and Y. Zhang, "Existence of pure-strategy equilibria in bayesian games: a sharpened necessity result," *International Journal of Game Theory*, vol. 46, no. 1, pp. 167–183, 2017.

[44] D. Fudenberg and J. Tirole, "Game theory," The MIT press, Tech. Rep., 1991.

[45] C. G. Lopes and A. H. Sayed, "Diffusion least-mean squares over adaptive networks: Formulation and performance analysis," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3122–3136, 2008.

[46] A. H. Sayed, S.-Y. Tu, J. Chen, X. Zhao, and Z. J. Towfic, "Diffusion strategies for adaptation and learning over networks: an examination of distributed strategies and network behavior," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 155–171, 2013.

[47] C. C. Aggarwal, "Outlier analysis," in *Data mining*. Springer, 2015, pp. 237–263.

[48] M. Kárnỳ, J. Kadlec, and E. L. Sutanto, "Quasi-bayes estimation applied to normal mixture," in *Preprints of the 3rd European IEEE Workshop on Computer-Intensive Methods in Control and Data Processing*. Praha, 1998, pp. 77–82.

[49] J. Fernandez-Bes, J. Arenas-Garcia, and A. H. Sayed, "Adjustment of combination weights over adaptive diffusion networks," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 6409–6413.

[50] L. A. Wolsey, "Mixed integer programming," *Wiley Encyclopedia of Computer Science and Engineering*, 2008.

[51] 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 (S'16)received his 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 distributed optimization with the application in smart power grids. Mr. Latifi is a student member of the IEEE.

Azam Khalili (M'12) 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.

Amir Rastegarnia (M'12) 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.

Saeid Sanei (SM'05) received his Ph.D. in signal processing from Imperial College London, U.K. in 1991. He has co-authored three monograms, a number of book chapters, and more than 350 papers in peer reviewed journals and conference proceedings. His research interests include adaptive filtering, cooperative learning, multiway, multimodal, and multichannel signal processing with applications to biomedical, audio, and communication signal processing, images, and biometrics. Dr Sanei has been an Associate Editor for the IEEE Signal Processing Letters, IEEE Signal Processing Magazine, Journal of Computational Intelligence and Neuroscience, and Elsevier Journal of Scientia Iranica. Currently, he is a Professor of Signal Processing and Machine Learning in Nottingham Trent University, Nottingham, U.K.