Economic Assessment for Battery Swapping Station Based Frequency Regulation Service

Xinan Wang, *Student Member, IEEE*, and Jianhui Wang, *Senior Member, IEEE*

***Abstract*—Battery swapping stations (BSSs) have great potential in providing fast frequency regulation service (FFRS) owing to their large battery storage capacity. However, compared to a regular BSS, a BSS providing FFRS faces the following financial risks: first, higher infrastructure investment to support the vehicle-to-grid services; second, higher battery aging costs due to FFRS; third, FFRS causes uncertainties to batteries' charging costs. Under such a context, in this article, we propose an economic risk assessment model for the BSS-based FFRS by comparing its economics with a regular BSS. In this model, the value at risk (VaR) of daily revenue and the long-term return on investment (ROI) of a regular BSS and a BSS providing FFRS are compared. The assessment results are obtained in three steps: first, we develop the operation and economic models for the BSS-based FFRS. Next, the mathematical models of the VAR and ROI analyses are formulated. Finally, the VaR of daily revenue is compared through a statistical analysis of a large number of scenarios. The ROI comparison is conducted by a policy gradient based reinforcement learning algorithm, which can handle the nonconvexity and stochastic dynamics brought by the electric vehicle visits. The practicality of the proposed framework is demonstrated by using the real ancillary market data from utilities and the traffic count data from onsite traffic sensors.**

***Index Terms*—Battery swapping station (BSS), fast frequency regulation service (FFRS), policy gradient (PG), reinforcement learning.**

NOMENCLATURE

Indices

n	Index referring to electric vehicle (EV).
t	Index referring to time horizon by hour.
j	Index referring to regulation area control error signal.

Parameters

$\overline{N}_{\text{bss}}$	Battery stock capacity of a battery swapping station (BSS).
S_{ip}	Initial state of charge (SOC) in the battery to be precharged.
S_{ic}	Initial SOC in the coming EV.
$\overline{S}_f/\underline{S}_f$	Battery SOC upper/lower bounds for participating in FFRS.
$\overline{P}_f/\underline{P}_f$	Battery regulation capacity upper/lower limits (kW).
P_r	Per battery regulation capacity (kW).
F_{ch}	Charging price paid by EV owners ($/kWh).
F_l	Real-time locational marginal price
$F_{\text{p/c}}$	Market performance/capacity clearance price
β	Ratio of BSS visit count to traffic flow.
η_b	Battery charge/discharge efficiency.
Q	Battery capacity of an EV
ϕ	Performance score.
λ	Mileage ratio.
ε	Battery value depreciation rate by year.
γ	Interest rate.
f	Price inflation rate.
$\sigma^{+/-}$	Fractional regulation up/down signal.
$\hat{q}$	Per unit hourly cumulative battery SOC gain due to FFRS.
v^F	Uncertainty band for EV visiting ratio.
T	Battery age.
T_{life}	Battery life in years.
NCY	Number of charge/discharge cycle per year.

Variables

N_{pc}	Number of precharged batteries.
N_{pf}	Number of batteries participating in FFRS.
N_{fc}	Number of batteries with SOC$>\overline{S}_f$ due to FFRS.
N_{od}	Number of batteries with SOC$<\underline{S}_f$ due to FFRS.
N_{red}	Number of redundant fully charged batteries.
N	Number of BSS visit count.
N'	Number of batteries precharged for EVs.
C	Cost of a BSS on purchasing energy from the grid.
C_B	Battery aging cost.
I	Revenue of a BSS on charging service.
P_c	Energy needed to compensate overdischarged batteries to 50% SOC.
P_p	Energy needed to precharge batteries.
P_f	Energy needed to fully charge the batteries with SOC$>\overline{S}_f$.
P_u	Energy needed to charge the replaced batteries to 50% SOC.
R	Deployed FFRS capacity from a BSS.
$\hat{R}$	Scheduled FFRS capacity from a BSS.
S_f	SOC in the battery participating in FFRS.
$S_{\text{io/ig}}$	Battery's SOC that is lower/higher than the bounds $[\overline{S}_f/\underline{S}_f]$ due to participating in FFRS.
S_{ip}	SOC in the battery for precharging.

Manuscript received October 25, 2019; revised February 3, 2020 and March 21, 2020; accepted April 1, 2020. Date of publication April 14, 2020; date of current version September 18, 2020. Paper 2019-AAAE-1320.R2, approved for publication in the IEEE TRANSACTION ON INDUSTRY APPLICATIONS by the Advanced Approaches and Applications for Electric Vehicle Charging Demand Management of the IEEE Industry Application Society. This work was supported by the U.S. Natural Science Foundation under Grant 1800716. *(Corresponding author: Jianhui Wang.)*

The authors are with the Department of Electrical and Computer Engineering, Southern Methodist University, Dallas, TX 75205 USA (e-mail: xinanw@smu.edu; jianhui@mail.smu.edu).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIA.2020.2986186

B Revenue of a BSS participating in FFRS.
B_{ch} Revenue of a BSS from charging service.
B_{fr} Revenue of a BSS from FFRS.
r^* Economic risk index for a BSS participating in FFRS.
a EV visiting ratio coefficient.
τ Decision trajectory of a.
P_θ Probability of the AI agent to choose action trajectoryτ.
θ Neural network parameters.

I. Introduction

With the rapid growth of renewable energy penetration in the U.S. [1], its inherent uncertainty and intermittency bring challenges to the stability of grid frequency [2], [3]. Therefore, the demand for the fast frequency regulation units surges. Compared with conventional regulation resources such as pumped hydro storage plants, and combustion turbines [4], fast ramping units, such as battery storage systems and flywheel energy storage systems, are significantly advanced in energy efficiency and accuracy to respond to an area control error (ACE) signal. With those advantages, a relatively small-scale deployment of fast ramping units can replace a large amount of existing generation-based regulation resources. According to a report of Pacific Northwest National Laboratory (PNNL), California Independent System Operator (CAISO) could reduce 40% of its regulation requirement if a proper mix of generation and storage-based regulation assets are deployed [5]. However, replacing a large scale of conventional regulation resources with those fast ramping unit is financially unaffordable [6] by utilities.

Inspired by the success of letting electric vehicle (EV) fleets participate in energy arbitrage [6], [8], we alternatively investigate the profitability of providing fast frequency regulation services (FFRSs) through the EV batteries. A typical EV battery has a capacity between 12 and 90 kWh [9]. Those distributed batteries can be managed through a proper battery charging strategy to offer FFRS to the grid. Currently, most of the EV-based FFRS strategies are developed for plug-in EV (PEV) [10]–[15]; very few studies focus on battery swappable EV (BSEV) based FFRS [16], [17] as battery swapping technology has only matured in recent years [18], [19]. In general, BSEV has the following advantages over PEV in providing FFRS: 1) a single battery swapping station (BSS) can reserve a large regulation capacity through the stored batteries to bid into the market, which requires at minimum 1 MW [20] to bid in, whereas PEVs are sparsely connected in the grid, they have to be centrally managed through PEV aggregators to concentrate sufficient capacity, and 2) a BSS can respond to the ACE signals in real time, while the hierarchical communication network between PEV aggregators and PEVs suffers from communication delay [13]–[15].

Although BSSs have been recognized as ideal FFRS resources [16], [17], they might also need additional hardware and software to support the ancillary services to the grid. A BSS that provides FFRS requires hybrid ac–dc/dc–ac inverters to support the bidirectional energy flow, while a regular BSS only requires cheaper ac–dc inverters. In addition, BSS-based FFRS relies on more complex software and robotic systems to properly allocate the batteries between FFRS and charging services. We summarize the financial risks for a BSS participating in FFRS into the following three aspects:

1) higher infrastructure investment to support the vehicle-to-grid services,
2) higher battery aging costs due to FFRS, and
3) FFRS causes uncertainties to batteries' charging costs.

Currently, there is no study to evaluate the economic risks of such a BSS-based FFRS business model, partially because there is no complete model for such a business. Besides, the long-term economic risks assessment for the BSS-based FFRS is a complicated Markov decision process (MDP), which involves the nonconvex dynamics caused by uncertain EV visits and is challenging to solve.

In this article, we model the BSSs as both energy consumers and ancillary service providers in the hourly ahead market, and its posterior economic assessment is conducted and compared with a regular BSS. The comparison metrics include value at risk (VaR) of daily revenue and long-term return on investment (ROI). The comparison of VaR of daily revenue is conducted through a statistical analysis on a large number of scenarios. The comparison of long-term ROI is formulated into a stochastic optimization problem that is solved using a policy gradient (PG) based reinforcement learning algorithm where the artificial intelligence (AI) agents learn to collect an increasingly higher reward by iteratively updating their action policies under different states given by the environment. In our framework, the environment is the BSS-based FFRS model; an action is the number of EV hourly visits to the BSS; a state refers to the operation status of the BSS; a reward shows the revenue ratio between a regular BSS and a BSS providing FFRS. The proposed method offers the following contributions.

1) *Complete modeling for the BSS-based FFRS*: the model simulates the continuous operation of a BSS providing FFRS. It includes the battery management strategy for FFRS and swapping service, battery aging model and EV visiting model.
2) *Reinforcement learning-based economic analysis for BSS-based FFRS*: the nonconvexity and stochastic dynamics in the problem are handled by an AI agent. To the best of authors' knowledge, this is the first article that proposes the methodology for long-term economic analysis of the BSS-based FFRS.
3) *Case studies using real-world data*: The practicality of the assessment results is demonstrated by using the real ancillary market data from utilities and the traffic count data from onsite traffic sensors.

The remainder of this article is organized as follows. Section II presents the model details of a BSS participating in FFRS. Section III introduces the PG-based AI agent formulation and its training process. Section IV presents the financial risk assessment of the proposed BSS-based FFRS model using real-world data. Section V provides some concluding remarks.

II. Battery Charging and Management Strategy

A. Battery Management Strategy for FFRS

To guarantee supply and demand balance between charged batteries and visiting EVs, a BSS requires every EV to submit a

battery swapping request online prior to its visit. Then, BSS can precharge batteries to full SOCs for visiting EVs on an hourly basis. The battery chargers considered in this study are the Level 3 dc chargers [21], which can fully charge a depleted battery within 30 min. Other batteries in BSS that are not scheduled for precharging provide FFRS to the grid. Since the locational marginal price of electricity and the FFRS market price are cleared on an hourly basis, we set the time step to be one hour as well in this study, which means the battery swapping request needs to be at least one hour earlier before the vehicle arrives.

The SOCs for the batteries participating in FFRS change following the ACE signals:

$$\hat{q}_t = \sum_{j=1}^{J} \left(\frac{\delta_{j,t}^{+}}{\eta_b} + \delta_{j,t}^{-} \cdot \eta_b \right) \cdot \Delta t \tag{1}$$

$$S_{f,t}^{n} = S_{f,t-1}^{n} - \frac{\hat{q}_t \cdot P_r}{U}, (\underline{S}_f \le S_{f,t}^{n} \le \overline{S}_f) \tag{2}$$

where $\delta_{j,t}^{+/-}(|\delta_{j,t}^{+/-}| \le 1)$ is the jth fractional frequency regulation signal (ACE signal) at the hour t. "+" denotes regulation-up, which requires the batteries to discharge, whereas "−" denotes regulation-down, which requires the batteries to be charged. The ACE signals used in this study are the RegD signals from PJM [22]. η_b is the battery charge/discharge efficiency. Δt represents the time interval between ACE signals. At each hour, a battery has a fractional energy gain as $\hat{q}_t \cdot P_r$ is the regulation capacity from each battery. The product of $\hat{q}_t$ and P_r is the energy gain for a battery at time t due to FFRS. In (2), U refers to the battery capacity and $\frac{\hat{q}_t \cdot P_r}{U}$ represents the SOC updates for batteries. $S_{f,t}^{n}$ indicates the SOC of the nth battery participating in FFRS at time t. If a battery's SOC is out of the bound $[\overline{S}_f / \underline{S}_f]$, that battery must quit FFRS in the next hour to ensure the BSS has a firm regulation capability on an hourly basis. The total regulation capacity available from the BSS at time t is $\hat{R}_t$ as shown as follows:

$$\hat{R}_t = N_{\mathrm{pf},t} \cdot P_r \tag{3}$$

$$N_{\mathrm{pf},t} = N_{\mathrm{pf},t-1} - N'_{t+1} + N_{t-1} - N_{\mathrm{od},t-1} - N_{\mathrm{fc},t-1} + N_{\mathrm{od}.t-2} (0 \le N_{\mathrm{pf},t} \le \overline{N}_{\mathrm{bss}}) \tag{4}$$

$$N'_{t+1} = H[N_{t+1} - N_{\mathrm{fc},t-1} - N_{\mathrm{red},t}] \cdot (N_{t+1} - N_{\mathrm{fc},t-1} - N_{\mathrm{red},t}) \tag{5}$$

$$N_{\mathrm{red},t} = H[N_{\mathrm{fc},t-2} + N_{\mathrm{red},t-1} - N_t] \cdot (N_{\mathrm{fc},t-2} + N_{\mathrm{red},t-1} - N_t). \tag{6}$$

$N_{\mathrm{pf},t}$ is the number of batteries participating in FFRS at the hour t. At each time interval of ACE signals, the energy required for FFRS will be evenly distributed to the $N_{\mathrm{pf},t}$ available batteries, and each battery will provide $\frac{\hat{R}_t \cdot \delta_{j,t}^{+/-}}{N_{\mathrm{pf},t}}$ kW capacity. In (4), N'_{t+1} denotes the number of cells that need to be precharged for the coming EVs at the next hour. N'_{t+1} equates to the number of coming EVs N_{t+1} minus $N_{\mathrm{fc},t-1}$ and $N_{\mathrm{red},t}$ as shown in (5), where $H(x)$ is a Heaviside step function, $H(x) = 1$ when $x > 0$, otherwise $H(x) = 0$. $N_{\mathrm{fc},t-1}$ refers to the number of batteries with SOC $> \overline{S}_f$ due to FFRS at the previous hour. Those batteries quit FFRS and are fully charged at time t to serve the visiting EVs at time $t+1$. $N_{\mathrm{red},t}$ is the number of the redundant fully charged batteries at the hour t. $N_{\mathrm{red},t}$ exists when: $N_{\mathrm{fc},t-2} + N_{\mathrm{red},t-1} > N_t$, which is shown in (6). In (4), N_{t-1} represents the batteries replaced from the visiting EVs at time $t-1$, they are charged to 50% SOC and then put into FFRS at time t because a 50% SOC provides a battery with equal ramping-up/down potential. $N_{\mathrm{od},t-1}$ is the number of batteries with SOC lower than $\underline{S}_f$ due to FFRS at time $t-1$. Those batteries are charged to 50% SOC at hour t and put back to FFRS at time $t+1$. The battery management logics shown from (3)–(6) are plotted in Fig. 1.

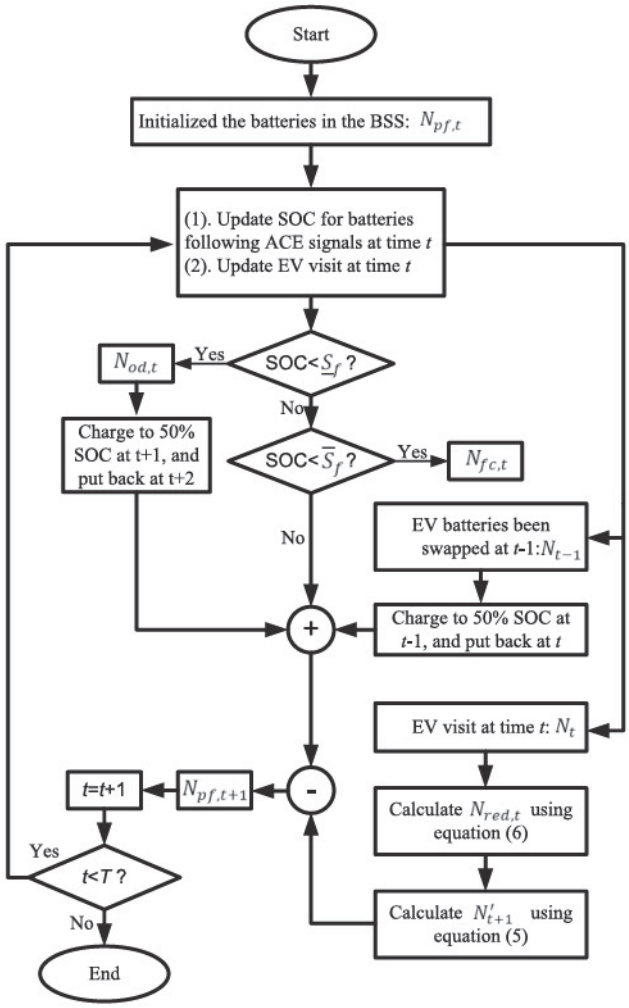

Fig. 1. Battery management logics for the BSS-based FFRS.

The income of BSS from FFRS is calculated using the PJM model as shown in the following equation, which is introduced in [23]

$$B_{\mathrm{fr},t} = \hat{R}_t \cdot \phi \cdot (\lambda_t F_{p,t} + F_{c,t}). \tag{7}$$

In (7), ϕ is presented as the performance score, which measures the accuracy of a BSS in following the ACE signals; λ_t refers to the mileage ratio at time t, which is the ratio of movement between the fast regulation signal and the regular

regulation signal in a given time period [24]; F_p is marked as the FFRS performance market clearing price; F_c denotes the FFRS capacity market clearing price; and $B_{\mathrm{fr},t}$ indicates the revenue that BSS receives from participating in FFRS.

B. Battery Management for Swapping Services

The other income source for the BSS is from battery swapping services. Equation (8) shows the process of selecting the precharging batteries, *sort*(A, n, *'descent'*) is a function that sorts the elements in the set A in descent and chooses the top n elements.$S^n_{\mathrm{ip},t}$is the initial SOC of the nth battery selected for precharging at time t. The batteries that are selected for precharging are those with the top N'_{t+1} highest SOCs in the BSS, so that the charging cost is minimized. In (9), $P_{p,t}$ refers to the amount of energy needed to precharge batteries. In (10), $P_{f,t}$ represents the energy required to fully charge the batteries with SOC $> \overline{S}_f$ due to FFRS as $S_{\mathrm{ig},t}(1 > S_{\mathrm{ig},t} > \overline{S}_f)$ indicating the initial SOC of those batteries. BSS also charges the batteries with SOCs $< \underline{S}_f$ due to FFRS and the batteries from the visiting EVs to 50% SOC, the energy needed is shown in (11). $S_{\mathrm{ic},t}$ indicates the initial SOC of the battery from a visiting EV $(0 < S_{\mathrm{ic},t} < \underline{S}_f)$ and $S_{\mathrm{io},t-1}$ refers to the initial SOC of a battery with SOC $< \underline{S}_f$ due to FFRS during the last hour

$$\boldsymbol{S}_{\mathrm{ip},t} = \mathrm{sort}(S_{f,t-1}, N'_{t+1}, '\mathrm{descent}') \tag{8}$$

$$P_{p,t} = \left(N'_{t+1} - \sum_{n=1}^{N'_{t+1}} S^n_{\mathrm{ip},t}\right) \cdot U \quad \forall S_{\mathrm{ip},t} \in \boldsymbol{S}_{\mathrm{ip},t} \tag{9}$$

$$P_{f,t} = \left(N_{\mathrm{fc},t-1} - \sum_{n=1}^{N_{\mathrm{fc},t-1}} S^n_{\mathrm{ig},t-1}\right) \cdot U \tag{10}$$

$$P_{c,t} = \frac{1}{2}\left(N_{\mathrm{od},t-1} + N_t - \sum_{n=1}^{N_{\mathrm{od},t-1}} S^n_{\mathrm{io},t-1} - \sum_{n=1}^{N_t} S^n_{\mathrm{ic},t}\right) \cdot U. \tag{11}$$

In (12), C_t shows the total charging cost for a BSS at time t and $F_{l,t}$ refers to the locational marginal price of electricity at time t. In (13), I_t indicates the income of the BSS received from providing battery swapping services to EVs at time t. F_{ch} is represented as the charging price in \$/kWh paid by customers. In (14), $B_{\mathrm{ch},t}$ is the revenue the BSS receives from charging services at time t, which equates to the differences between I_t and C_t

$$C_t = F_{l,t} \cdot (P_{p,t} + P_{f,t} + P_{c,t}) \tag{12}$$

$$I_t = \sum_{n=1}^{N_t} F_{\mathrm{ch}} \cdot U \cdot (1 - S^n_{\mathrm{ic},t}) \tag{13}$$

$$B_{\mathrm{ch},t} = I_t - C_t. \tag{14}$$

C. Battery Aging Model

Battery aging cost is included in our analysis because FFRS induces higher cycle aging costs to batteries as they follow the ACE signals. For instance, it is considered as one charge/discharge cycle if a battery's SOC falls below $\underline{S}_f$ while providing FFRS. We adopt a widely used depth of discharge (DOD) based battery cycle life model [28] into our analysis, as shown as follows:

$$T_{\mathrm{life}} = \frac{n\mathrm{Life}(\mathrm{DOD})}{\mathrm{NCY}} \tag{15}$$

where T_{life} denotes the battery life in years, which is obtained through dividing the battery cycle life $n\mathrm{Life}(\mathrm{DOD})$by the number of operation cycles per year (NCY). The battery cycle life $n\mathrm{Life}$ is a function of DOD as introduced in [28]. Then, we model the per-cycle aging cost of a battery by evenly distributing its annual value depreciation to each of its charge/discharge cycles in that year

$$C_{B,t} = \sum_{n=1}^{N_{\mathrm{od},t}+N_t} \frac{U \cdot F_B \cdot (1-\varepsilon)^{T_n} \cdot \varepsilon}{\mathrm{NCY}}, (0 \le T_n \le T_{\mathrm{life}}). \tag{16}$$

Equation (16) shows the total battery aging cost of a BSS participating in FFRS at time t, which includes the batteries from the visiting EVs and the batteries that quit the FFRS due to low SOCs. $F_B$refers to the price of battery in \$/kWh; εindicates the value depreciation rate of a battery in a year. $U \cdot F_B \cdot (1-\varepsilon)^{T_n}$ denotes the current value of the nth battery at its age T_n. The value depreciation of this battery for the current year is represented as $U \cdot F_B \cdot (1-\varepsilon)^{T_n} \cdot \varepsilon$. Thus, the per-cycle aging cost of this battery is derived by dividing its value depreciation for the current age by its operation cycle per year NCY. In our model, we consider the age distribution of the batteries in a BSS is random between 0 to T_{life}. The battery aging cost for a regular BSS is formulated in (17), which only considers the batteries aging due to powering EVs. In a regular BSS, batteries have a longer life, a smaller NCY, and a lower value depreciation rate compared with the batteries in a BSS that participate in FFRS due to the lower use intensity. We use T^*_{life}, NCY^*,and ε^* to indicate the length of battery life, number of cycles per year and value depreciation rate for those batteries in a regular BSS, respectively

$$C^*_{B,t} = \sum_{n=1}^{N_t} \frac{U \cdot F_B \cdot (1-\varepsilon^*)^{T_n} \cdot \varepsilon^*}{\mathrm{NCY}^*}, (0 \le T_n \le T^*_{\mathrm{life}}). \tag{17}$$

D. Model the EV Visit Uncertainty

Many research works model the EV uncertainty over time using specific distributions, such as normal distribution, and Poisson distribution [25]–[27]. These models are accurate when being applied to a large number of EVs but not for relatively small EV traffics experienced by a BSS. Since the service model of a BSS is similar to a gas station, we assume the BSS visit pattern is the same as the gas station visit pattern. GasBuddy [29] conducts a statistic analysis on more than 32.6 million customer trips to gas stations in 2018 and generates a 24-hour visit percentage chart (ratio between each hour's visits and the daily visits) of gas stations, which is plotted in red in Fig. 2(a). The blue curve in Fig. 2(a) shows the 24-hour average traffic flow (TF) of 120 days in percentage from the road I-280 in San Jose,

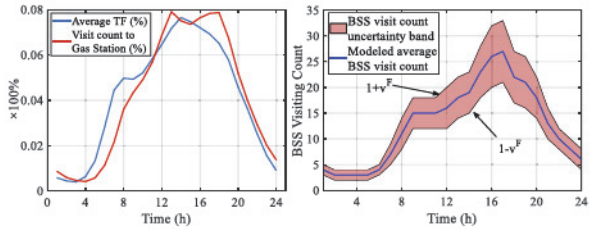

Fig. 2. (a) TF count versus gas station visit. (b) BSS visit uncertainty band.

California [30] in 2017. These two curves are highly similar to one another. We also discover this similarity in other traffic data that we collected. Therefore, we assume the BSS visit count is linearly related to TF data

$$N_t = \beta N_{f,t}. \tag{18}$$

In (18), β denotes the EV visiting ratio and $N_{f,t}$ refers to the TF at time t. The value of β can be adjusted according to the EV penetrations. However, if we consider the EV visits of each day as a unique pattern, then an uncertainty band is needed to simulate all the EV visits that belong to this pattern. Therefore, we apply an uncertainty band v^F on N_t, which can be found in (19). The value of v^F should be set to preserve the pattern feature. In this optimization problem, N_t is an integer variable to be decided at each hour by the agent within the uncertainty band formed by $(1 - v^F)\beta N_{f,t}$ and $(1 + v^F)\beta N_{f,t}$, as shown in Fig. 2(b)

$$(1 - v^F)\beta N_{f,t} \le N_t \le (1 + v^F)\beta N_{f,t}. \tag{19}$$

E. Metrics for the Economic Assessment

The following equation shows the revenue received by a BSS participating in FFRS from t_0 to t_e

$$B = \sum_{t=t_0}^{t_e} (B_{\mathrm{ch},t} + B_{\mathrm{fr},t} - C_{B,t}). \tag{20}$$

The following equation shows the revenue model of a regular BSS is presented, which only includes the incomes from battery swapping services:

$$B^* = \sum_{t=t_0}^{t_e} \left[\sum_{n=1}^{N_t} (F_{\mathrm{ch}} - F_{l,t}) \cdot (U - S^n_{\mathrm{ic},t}) - C^*_{B,t} \right]. \tag{21}$$

In this article, we introduce two metrics to compare the economics between a BSS providing FFRS and a regular BSS, they are the VaR of daily revenue and the long-term ROI ratio. VaR of daily revenue directly reflects the short-term cash flows of both business models whereas ROI reflects the long-term profitability of an investment. For an investor who owns the BSS-based FFRS business for P subperiods, the total profit gain is the net present value (NPV) [31] of its subperiod cumulative revenues, which is shown as follows:

$$\mathrm{NPV} = \sum_{p=1}^{P} \frac{B_p \cdot (1+f)^p}{(1+\gamma)^p} \tag{22}$$

where f represents the price inflation rate, γ indicates the interest rate, and p refers to the number of subperiods. The long-term ROIs of a BSS providing FFRS and a regular BSS are $\mathrm{ROI} = \frac{\mathrm{NPV}}{\mathrm{IV}}$ and $\mathrm{ROI}^* = \frac{\mathrm{NPV}^*}{\mathrm{IV}^*}$, respectively, where IV indicates the investment for a BSS providing FFRS and IV* refers to the investment for a regular BSS. The investors for BSSs expect a higher long-term ROI by letting the BSSs participate in FFRS. From (23), we can prove that if $\max(\frac{B^*_p}{B_p}) \le \frac{\mathrm{IV}^*}{\mathrm{IV}}$ is satisfied at each subperiod, then $\frac{\mathrm{NPV}}{\mathrm{IV}} \ge \frac{\mathrm{NPV}^*}{\mathrm{IV}^*}$ (or $\mathrm{ROI} \ge \mathrm{ROI}^*$). First, let us transform $\frac{\mathrm{NPV}}{\mathrm{IV}} \ge \frac{\mathrm{NPV}^*}{\mathrm{IV}^*}$ into $\frac{\mathrm{NPV}}{\mathrm{NPV}^*} \le \frac{\mathrm{IV}^*}{\mathrm{IV}}$. Then, $\frac{\mathrm{NPV}^*}{\mathrm{NPV}}$ can be expanded using (22) as

$$\frac{\mathrm{NPV}^*}{\mathrm{NPV}} = \frac{\sum_{p=1}^{P} \frac{B^*_p \cdot (1+f)^p}{(1+\gamma)^p}}{\sum_{p=1}^{P} \frac{B_p \cdot (1+f)^p}{(1+\gamma)^p}}. \tag{23}$$

Define $r_p = \max(\frac{B^*_p}{B_p}), (1 \le p \le P)$, where r_p indicates the highest revenue ratio for a BSS providing FFRS compared with a regular BSS at subperiod p. If $r_p \le \frac{\mathrm{IV}^*}{\mathrm{IV}}$ is satisfied at each subperiod, we fit $\frac{\mathrm{IV}^*}{\mathrm{IV}}$ into (23) to have the following equation:

$$\frac{\sum_{p=1}^{P} \frac{B^*_p \cdot (1+f)^p}{(1+\gamma)^p}}{\sum_{p=1}^{P} \frac{B_p \cdot (1+f)^p}{(1+\gamma)^p}} \le \frac{\sum_{p=1}^{P} \frac{\mathrm{IV}^* \cdot B_p \cdot (1+f)^p}{\mathrm{IV} \cdot (1+\gamma)^p}}{\sum_{p=1}^{P} \frac{B_p \cdot (1+f)^p}{(1+\gamma)^p}} = \frac{\mathrm{IV}^*}{\mathrm{IV}}. \tag{24}$$

Combine (23) and (24), we can conclude $\frac{\mathrm{NPV}}{\mathrm{IV}} \ge \frac{\mathrm{NPV}^*}{\mathrm{IV}^*}$ or $\mathrm{ROI} \ge \mathrm{ROI}^*$ when $r_p \le \frac{\mathrm{IV}^*}{\mathrm{IV}}, \forall p$. Therefore, we use r_p to quantitatively represent the investment risk. A higher r_p presents a higher financial risk for a BSS participating in FFRS.

To evaluate the financial risk of the BSS-based FFRS, we design an optimization problem as: $r_p = \max(\frac{B^*_p}{B_p})$. The inputs to this optimization problem are $\delta^{+/-}_{j,t}$, $F_{l/\mathrm{p/c},t}$, λ_t, and ϕ, which are the ancillary service market data released by utilitiy [22]. The uncertain variable in this problem is the EV visit count N_t. The stochastic parameter is the intial SOC of the battery in each coming EV.

The VaRs of daily revenue between the two business models are compared through a statistical analysis on different scenarios. In each scenario, the daily EV visit pattern is uniform random sampled within the boundaries defined by (19). Given that the actual EV visit pattern is decided by the customer behaviors, and it can be any curve within the uncertainty band. When comparing the ROIs of the two business models, an AI agent is trained to find an EV visit pattern $[N_1, N_2, N_3, \ldots, N_{t_e}]$ within the uncertainty band that yields the $\max(\frac{B^*_p}{B_p})$. The value of N_t at time t impacts the operation state trajectory of BSS and the AI agent's decision trajectory thereafter. Such a decision-making chain turns the problem into a MDP. In this MDP, the state transition probability is defined as $P(s_{t+1}|s_t, a_t)$, where s_{t+1} and s_t refer to the operation states of BSS at time $t+1$ and t; a_t is the decision made at time t. All the state transition possibilities

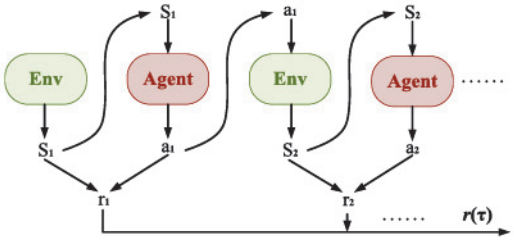

Fig. 3. State and action trajectory in a PG problem.

sum up as 1: $\sum_1^{s_N} P(s_{t+1}|s_t, a_t) = 1$, where s_N is the number of possible state transitions.

Because the EV visit uncertainty band is defined by percentage using v^F, we design a BSS visiting ratio coefficient $1+a_t$ to model the N_t uncertainty at each hour, which can be found in (25). $a_t$has a uniform variation boundary at each time step represented as $\pm v^F$. In this way, selecting value for a_t becomes the action the agent needs to decide and take at each time step

$$N_t = [(1 + a_t)\beta N_{f,t}], (-v^F \leq a_t \leq v^F). \tag{25}$$

III. PG-Based AI Agent

A. Introduction to PG

PG is one of the most popular reinforcement learning techniques, which applies gradient descent or ascent to its decision-making policy to optimize the expected long-trajectory cumulative reward in a dynamic process. PG outperforms many other traditional reinforcement learning approaches because PG does not suffer from problems such as the intractability problem resulting from uncertain state information and the complexity arising from contiguous states and actions [32]. Therefore, it is suitable to solve the problem in (23), which also contains long-chain nonconvex dynamics. In each training batch, the agent plays multiple episodes and makes decisions on a_t at each state s_t based on its current policy $\pi(s, a, \theta) = P_\theta\{a_t = a|s_t = s, \theta\}$. We denote τas a full state-action trajectory $\tau = \{s_1, a_1, s_2, a_2, \ldots, s_{t_e}, a_{t_e}\}$. Equation (26) shows $\boldsymbol{P}_\theta(\tau)$ is the possibility for the AI agent to travel through this trajectory, which is the chain-product of the state transition possibility$P(s_{n+1}|s_n, a_n)$ and the decision possibility $P_\theta(a_n|s_n)$. The environment decides the state transition possibility P and the AI agent determines the decision possibility P_θ, this process is shown in Fig. 3. The PG-based reinforcement learning process trains an AI agent to optimize its decision-making policy parameters θso that the possibility of choosing the trajectory that yields the highest rewards r is maximized

$$\boldsymbol{P}_\theta(\tau) = P(s_1)P_\theta(a_1|s_1)P(s_2|s_1, a_1)P_\theta(a_2|s_2)\cdots. \tag{26}$$

Before starting the next batch of training, the AI agent updates its policy parameter θ based on the current batch's results. The PG-based reinforcement learning assumes the policy π is differentiable with respect to its parameterθ, which means $\frac{\partial\pi(s,a)}{\partial\theta}$ exists. Furthermore, PG broadcasts the reward that the agent receives from the current batch in the backpropagation-based policy parameter updating process as in the way shown in (27) and (28). The agent iteratively updates its policy parameters until the expected reward $\overline{r}_p$converges. A positive reward will encourage the current policy whereas a negative reward will punish it

$$\overline{r}_p = \sum_\tau r_p(\tau)\boldsymbol{P}_\theta(\tau) = E_{\tau\sim\boldsymbol{P}_\theta(\tau)}[r_p(\tau)]. \tag{27}$$

By taking partial derivative of $\overline{r}_p$ toward θ, the policy parameter updating coefficient $\Delta\theta$is formulated as follows [33]:

$$\begin{aligned}
\nabla\overline{r}_p &= \nabla E_{\tau\sim\boldsymbol{P}_\theta(\tau)}[r_p(\tau)] \\
&= \sum_\tau r_p(\tau)\nabla\boldsymbol{P}_\theta(\tau) + \boldsymbol{P}_\theta(\tau)\nabla r_p(\tau) \\
&\approx \sum_\tau r_p(\tau)\nabla\boldsymbol{P}_\theta(\tau) \\
&= \sum_\tau r_p(\tau)\boldsymbol{P}_\theta(\tau)\frac{\nabla\boldsymbol{P}_\theta(\tau)}{\boldsymbol{P}_\theta(\tau)} \\
&= \sum_\tau r_p(\tau)\boldsymbol{P}_\theta(\tau)\nabla\ln\boldsymbol{P}_\theta(\tau) \\
&= E_{\tau\sim\boldsymbol{P}_\theta(\tau)}[r_p(\tau)\nabla\ln\boldsymbol{P}_\theta(\tau)] \\
&\approx \frac{1}{K}\sum_{k=1}^{K} r_p(\tau^k)\nabla\ln\boldsymbol{P}_\theta(\tau^k) \\
&= \frac{1}{K}\sum_{k=1}^{K}\sum_{t=1}^{t_e} r_p(\tau^k)\nabla\ln P_\theta(a_t^k|s_t^k).
\end{aligned} \tag{28}$$

In the second step of (28), we ignore the term $\boldsymbol{P}_\theta(\tau)\nabla r_p(\tau)$; because the reward $r_p(\tau)$ is obtained through the interactions between actions $a(\tau)$ and the environment. In reality, the environment can be highly nonlinear or non-differentiable. Therefore, it is tricky to get the gradient of $r_p(\tau)$. Fortunately, in [34], the authors prove that we still can guarantee policy improvement and reach a true local optimum even we ignore the gradient of $r_p(\tau)$. In the fifth step, we apply a gradient trick [32] to replace $\frac{\nabla\boldsymbol{P}_\theta(\tau)}{\boldsymbol{P}_\theta(\tau)}$ with $\nabla\ln\boldsymbol{P}_\theta(\tau)$ ($\frac{\nabla\boldsymbol{P}_\theta(\tau)}{\boldsymbol{P}_\theta(\tau)} \approx \nabla\ln\boldsymbol{P}_\theta(\tau)$). This gradient approximation significantly reduces the computational cost of the algorithm. Because the form of $\boldsymbol{P}_\theta(\tau)$is usually considered as Gaussian distribution, it can be costly for the computer to calculate the gradient of its probability density function. However, the gradient of $\nabla\ln\boldsymbol{P}_\theta(\tau)$is much easier to calculate as the nonlinear formation is converted into a linear structure. We approximate the expectation $\nabla E_{\tau\sim\boldsymbol{P}_\theta(\tau)}[r_p(\tau)\nabla\ln\boldsymbol{P}_\theta(\tau)]$ with an empirical average $\frac{1}{K}\sum_{k=1}^{K} r(\tau^k)\nabla\ln\boldsymbol{P}_\theta(\tau^k)$ for the stochastic policy $\pi(s, a, \theta)$updates in the seventh step of (28). This approximation is widely adopted in PG methods [35]. The policy parameter θ updates are shown in the following equation:

$$\theta = \theta + \varepsilon\cdot\left(\frac{1}{K}\sum_{k=1}^{K}(r(\tau^k) - A\right)\nabla\ln\boldsymbol{P}_\theta(\tau^k) \tag{29}$$

ε is a deterministic learning rate and A is a baseline reward.

This baseline A is needed because the reward is constantly positive in our application. If $\nabla\overline{r}_p$is directly used in (28) to update θ, the policy $\pi(s, a, \theta)$will always be encouraged in

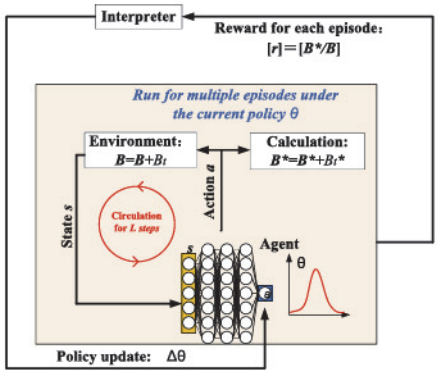

Fig. 4. AI agent training process.

any circumstances. We need to subtract a baseline from the received reward to ensure only the right policy is encouraged. This baseline is created by taking the mean reward of multiple random episodes that run prior to the agent training.

B. AI Agent Training Process

Equations (1)–(23) represent the training environment for the AI agent. This environment can receive action a_t from the AI agent and respond to the agent with a new operation state s_{t+1}. The dynamics of this environment are characterized by state transition probabilities $\mathcal{P}^a_{ss'} = P\{s_{t+1} = s' | s_t = s, a_t = a\}$, which is shown in Fig. 3.

The AI agent makes decisions a_t based upon the current state s_t. In the environment that we designed, the state vector is defined as $s_t = [N_{\mathrm{pf},t}, N_{\mathrm{fc},t-1}, N_{\mathrm{od},t-1}, N'_{t+1}, N_{t+1}]$because those variables are closely related to the objective r_p through (1)–(23): $B_t = \Phi(N_{\mathrm{pf},t}, N_{\mathrm{fc},t-1}, N_{\mathrm{od},t-1}, N'_{t+1}, N_{t+1})$, $B^*_t = \Psi(N_t)$. Φ denotes the process of calculating the temporal revenue of a BSS participating in FFRS using (1)–(20); Ψrepresents the process of calculating the temporal revenue of a regular BSS using (21). The agent training process is shown in Fig. 4; the pseudocode for the training process is in Algorithm I. A memory buffer M is predefined to store the historical state and action in each training batch; the policy updates using the data sampled from M after each batch of running.

IV. Case Studies

We use real-world data in the training environment, the environment specifications can be found in Table I. The ACE signals used in the case study are the RegD signals from PJM [22], the FFRS market data can be accessed from PJM [22], and the traffic count data in California can be obtained from PeMS [30]. The hourly market and traffic data for 2017 and 2018 is used in the case study. The initial number of batteries stored in BSS is randomly selected between 80 and 100 (the maximum battery storage capacity $\overline{N}_{\mathrm{bss}}$ is 200). The initial SOCs of the batteries in the BSS are set randomly between 0.2 and 0.8; the initial SOCs of visiting EVs are set randomly between 0.2 and 0.3; the uncertainty level of EV visiting flow v^F is set to be 30%; according to the current level of technology, the cycle life of lithium-ion batteries is between 2500 and 6000 cycles [28], [36] at 80% DOD. In this study, we consider the battery life can last 3500–4000 cycles. With a 30 kWh battery, the EV owners are expected to swap their battery on a daily basis (NCY* = 365) under a heavy use intensity. Therefore, the battery life T^*_{life} is over 10 years, and we conservatively set $T^*_{\mathrm{life}} = 10$ years. Since there is no theoretical study to investigate FFR S's influences on battery life, we can only estimate the NCY based on our simulation platform. Our analysis shows that batteries in BSS that participate in FFRS will experience one more charge/discharge cycle compare with those batteries that

Algorithm I: PG-Based AI Agent Training.

Input: FFRS market price $F_{p/c}$, locational marginal price F_l, per battery regulation limit P_r, traffic flow N_f, β and v^F, trajectory length L, batch size W, batch number V
Output: r_p
Initialize neural network *nn* and memory buffer M
Random run n episodes and get baseline A.
While batch number< V **and** r is not converged **do**
  Clear buffer M
  Initialize episode reward r_p
  For each episode in a batch **do**
    While t<L **do**
      $a_t \sim P_\theta(a_t|s_t)$
      $s_{t+1} \sim P(s_{t+1}|s_t, a_t)$
      $B_t =$
      $\Phi(N_{pf,t}, N_{fc,t-1}, N_{od,t-1}, N'_{t+1}, N_{t+1})$
      $B^*_t = \Psi(N_t)$
      $M \leftarrow M \cup [a_t, s_t]$
      $B = B + B_t$
      $B^* = B^* + B^*_t$
      $t = t+1$
    End While
    $r_p \leftarrow (r_p \cup \frac{B^*}{B})$
  End for
  $\Delta\theta \leftarrow \Gamma(P_\theta, \overline{r}, A)$ # Γrefers to the function of gradient calculation
  $\{s_M, a_M\} \leftarrow sample(M)$ # sample state and action pairs from buffer M
  $\theta \leftarrow \theta + \varsigma \cdot \Delta\theta$ #update the policy
End While

TABLE I
System Parameter Setting

Parameter	Value	Parameter	Value
η_b	0.9 [37]	F_l	PJM [22]
$\overline{N}_{bss}$	200	F_p	PJM [22]
U	30kWh [9]	F_c	PJM [22]
φ	0.98 [38]	λ	PJM [22]
TF data	PeMS [30]	φ	PJM [22]
β	0.05	$\delta^{+/-}$	PJM [22]
ε	0.01	P_r	15 kW
$\overline{S_f}/\underline{S_f}$	0.2/0.8	F_{ch}	$0.12/kWh
V	300	W	50
M	10e6	L	Hours of a month
v^F	30%	$\overline{N}_{bss}$	200
F_B	280 [39]	ϵ	0.2
T_{life}/T^*_{life}	5/10 [28]	NCY/NCY*	730/365 [28]

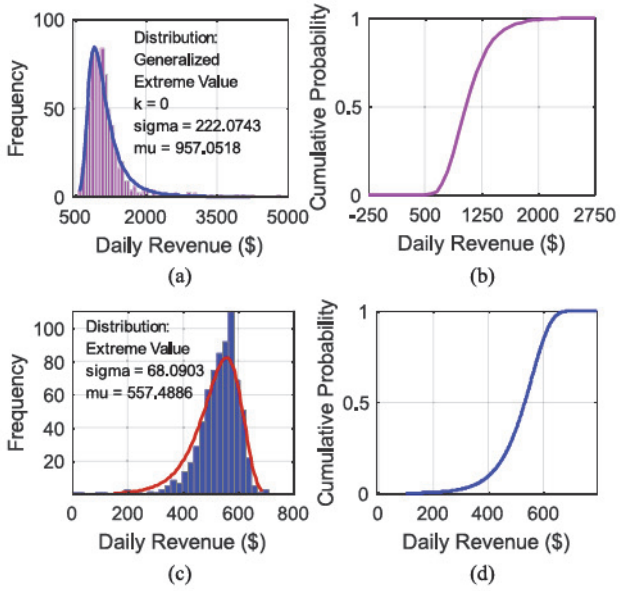

Fig. 5. (a) Daily revenue distribution and the associated PDF plot for a BSS participating in FFRS. (b) CDF of the distribution. (a) and (c) Daily revenue distribution and the associated PDF plot for a regular BSS. (d) CDF of the distribution in (c).

do not participate in FFRS or NCY = 730. Therefore, according to [28] and (15), under the same DOD level, we set battery life as $T_{\text{life}} = 5$ years.

The maximum training batch for the AI agent is $V = 300$. There are 50 episodes (W) in each batch, the decision trajectory length L is either 672, 720, or 744 depending on the number of hours within that month.

A. VaR of Daily Revenue Comparison

In this section, we compare the VaR of daily revenue between a BSS participating in FFRS and a regular BSS using the 730 day's data we collected. The EV visit pattern for each scenario is generated by adding uncertainties within the uncertainty band $\pm v^F$ to daily traffic data. The daily revenue distributions of the two business models during 2017 and 2018 are plotted in Fig. 5. For a BSS participating in FFRS, its daily revenue is close to a generalized extreme value distribution with $k = 0$, $\mu = 957.0518$, and $\sigma = 222.0743$ according to a data fitting analysis. The associated distribution and probability density function are plotted in Fig. 5(a) and the cumulative probability density is plotted in Fig. 5(b). For a regular BSS, the daily revenue follows an extreme value distribution with the $\mu = 557.4886$ and $\sigma = 68.0903$. The associated distribution and probability density functions are plotted in Fig. 5(c) and the cumulative probability density is plotted in Fig. 5(d).

According to both business models' daily revenue distributions, we calculate the corresponding VaR of daily revenue and summarize the results in Table II. For a regular BSS, the probability of having a daily revenue higher than \$600 is merely 17.08%. However, the daily revenue of a BSS participating in FFRS has a 99.33% probability to exceed \$600. A regular BSS has zero possibility to have a daily revenue higher than \$1000, while that probability is 55.92% for a BSS participating in FFRS.

TABLE II
VALUE AT RISK OF DAILY REVENUE COMPARISON

Daily revenue (\$)	Possibility of B	Possibility of B^*
≥200	1.0000	0.9933
≥400	1.0000	0.8967
≥600	0.9933	0.1708
≥800	0.8678	0.0000
≥1,000	0.5592	0.0000
≥1,200	0.2822	0.0000
≥1,400	0.1256	0.0000

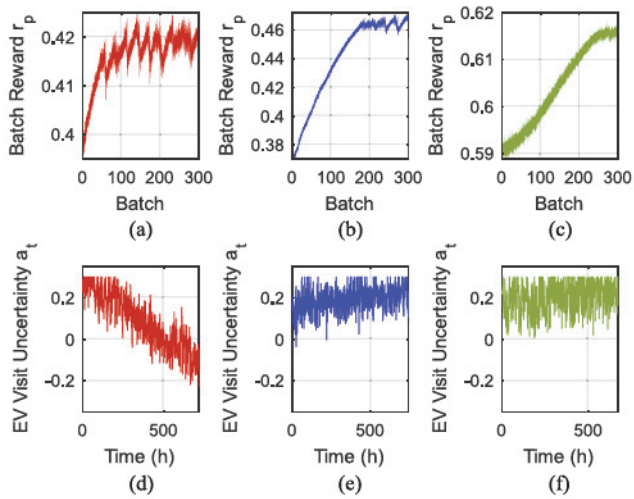

Fig. 6. (a) PG agent's batch reward band for April 2017. (b) PG agent's batch reward band for May 2017. (c) PG agent's batch reward band for February 2018. (d) EV visit uncertainty pattern for April 2017 that yields the r_p. (e) EV visit uncertainty pattern for May 2017 that yields the r_p. (f) EV visit uncertainty pattern for February 2018 that yields the r_p.

B. ROI Comparison Between the Two Business Models

To compare the long-term ROIs of the two business models, the risk index r_p needs to be calculated at each subperiod. Therefore, we use one month as the period base forr_p. As discussed in Section II-B, the EV visit pattern contains significant uncertainties. Different EV visit patterns result in various revenue for both the business models. For each month of 2017 and 2018, we train an AI agent to explore an EV visit pattern that causes the highest revenue risk r_p. The training process for April 2017, May 2017, and February 2018 are shown in Fig. 6.

The agent policy updates at each batch of training, the highest and the lowest reward at each batch are plotted as the upper and lower boundaries of the reward bands as shown in Fig. 6(a)–(c). For April 2017, the reward converges at around 0.42 after 100 batches and the highest reward is r_p= 0.4248 as shown in Fig. 6(a), which means the monthly revenue of the BSS providing FFRS is at least $\frac{B_p}{B_p^*} \geq \frac{1}{0.4248} = 2.3540$ times of the regular BSS in April 2017. Fig. 6(d) shows the EV visit uncertainty pattern in April 2017 that yields the r_p. The same training results for May 2017 and February 2018 are shown in Fig. 6(b) and (e) and (c) and (f), respectively. The values of r_p at each month of 2017 and 2018 are summarized in Table III.

In December 2018, the BSS providing FFRS earns at least 1040.25% more profits than a regular BSS. This is because in that month, the average FFRS capacity clearance price is 4.46 times higher than the other months. So we treat the r_p in

TABLE III
REVENUE RATIO SUMMARY FOR 2017 AND 2018

Month Of 2017	r_p	Revenue Increase	Month Of 2018	r_p	Revenue Increase
Jan	0.5749	73.94%	Jan	0.5361	86.54%
Feb	0.5859	70.69%	Feb	0.6179	61.85%
Mar	0.4689	113.27%	Mar	0.5474	82.63%
Apr	0.4248	135.40%	Apr	0.4187	138.83%
May	0.4721	111.82%	May	0.5021	99.17%
Jun	0.4778	109.29%	Jun	0.4672	114.06%
Jul	0.5166	93.57%	Jul	0.5414	84.70%
Aug	0.4463	124.06%	Aug	0.3547	181.91%
Sep	0.5234	91.07%	Sep	0.3851	159.66%
Oct	0.4157	140.58%	Oct	0.4455	124.47%
Nov	0.4453	124.55%	Nov	0.5066	97.40%
Dec	0.3826	161.73%	Dec	0.0877	1040.25%

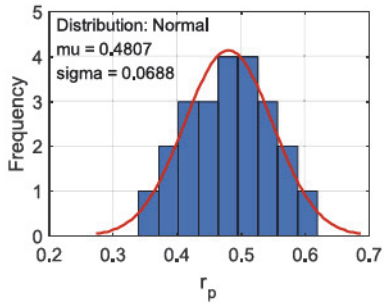

Fig. 7. Distribution of r_p.

TABLE IV
PROBABILITY CHART FOR ROI $\geq$ ROI*

$\frac{IV}{IV^*}$	$P(r_p \leq \frac{IV^*}{IV})$	$\frac{IV}{IV^*}$	$P(r_p \leq \frac{IV^*}{IV})$	$\frac{IV}{IV^*}$	$P(r_p \leq \frac{IV^*}{IV})$
1.0	1.0000	1.3	1.0000	1.6	0.9820
1.1	1.0000	1.4	0.9997	1.7	0.9409
1.2	1.0000	1.5	0.9966	1.8	0.8616

December 2018 as an outlier and fit the r_p of other months into an normal distribution as shown in Fig. 7, with its mean value μ as 0.4807 and standard deviation σ as 0.0688. Fig. 7 represents the distribution of r_p.

As discussed earlier, if $r_p < \frac{\text{IV}^*}{\text{IV}}$ is satisfied at each subperiod, then it is sufficient to guarantee ROI $\geq$ ROI*. According to the distribution of r_p, we calculate the confidence for a BSS participating in FFRS to have a higher monthly ROI than a regular BSS with respect to different investment ratios, which is shown in Table IV. The following statement can be made from the results: if the investment of a BSS participating in FFRS is no more than 30% higher of a regular BSS, we have 100% confidence to guarantee ROI $\geq$ ROI* no matter how long the investors hold this business. If the investment is 40% higher and the holding-period is 5 years, we have more than $0.9997^{60} = 98.37\%$ confidence that ROI $\geq$ ROI*. The risk of having a 5-year ROI $\geq$ ROI* will significantly increase if the investment for a BSS providing FFRS is 50% higher than a regular BSS.

C. Algorithm Performance Comparison

The proposed PG-based optimization framework does not require the explicit formulation of the problem and can avoid the insolvability and intractability of the traditional methods. It shows a good convergence toward the designed nonlinear dynamic optimization problem. To evaluate the quality of the solutions given by the PG-based algorithm, we compare the solutions with the results generated from the Monte Carlo simulation. For each month, 10 000 EV visit scenarios are uniformly sampled within the uncertainty band. The highest reward among the 10 000 scenarios is then compared with r_p. The comparison results are shown in Fig. 8.

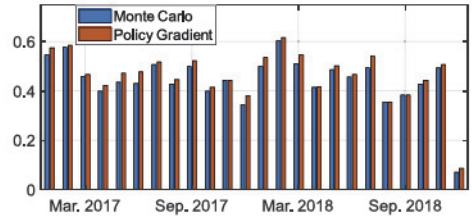

Fig. 8. Algorithm performance comparison with Monte Carlo.

From Fig. 8, it is seen that the result of the proposed PG-based algorithm is always better than the result of the Monte Carlo simulation. Therefore, the quality of solutions given by the proposed method in this article is guaranteed and can serve as a benchmark method for other relative studies.

V. CONCLUSION

In this article, we designed the operation and economic models for a BSS participating in FFRS and compared its economics with a regular BSS. The PG-based AI agent was used to deal with the nonconvex dynamics contained in the analysis. The results of this study demonstrated two important conclusions: 1) the proposed PG-based algorithm was capable of providing a high-quality solution to the economic analysis of BSS-based FFRS, and 2) with a proper system setup, a BSS participating in FFRS could have much higher ROI and revenue than a regular BSS. As technology improves, the profitability of the BSS-based FFRS can be more attractive in the future as the batteries' cost will be lower but life will be longer.

REFERENCES

[1] M. Hand *et al.*, "Renewable electricity futures study," National Renewable Energy Laboratory (NREL), 2013. [Online]. Available: https://www.nrel.gov/docs/fy12osti/52409-1.pdf. Accessed: Mar. 18, 2020.
[2] N. Nguyen and J. Mitra, "An analysis of the effects and dependency of wind power penetration on system frequency regulation," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 354–363, Jan. 2016.
[3] Y. Zhang, J. Wang, and Z. Li, "Uncertainty modeling of distributed energy resources: Techniques and challenges," *Current Sustain. Energy Rep.*, vol. 6, no. 2, pp. 42–51, Jun. 2019.
[4] L. Thomas, "Energy storage in PJM exploring frequency regulation market transformation," Kleinman Center for Energy Policy, University of Pennsylvania, Philadelphia, PA, Jul. 2017. [Online]. Available: https://kleinmanenergy.upenn.edu/sites/default/files/proceedingsreports/Energy%20Storage%20in%20PJM_0.pdf. Accessed: May 27, 2020.
[5] Y. V. Makarov, J. Ma, S. Lu, and T. B. Nguyen, "Assessing the value of regulation resources based on their time response characteristics," Pacific Northwest National Laboratory (PNNL), Jun. 2008. [Online]. Available: https://www.pnnl.gov/main/publications/external/technical_reports/PNNL-17632.pdf. Accessed: Mar. 18, 2020.

[6] International Renewable Energy Agency, "Electricity storage and renewables: Costs and markets to 2030," Oct. 2017. [Online]. Available: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2017/Oct/IRENA_Electricity_Storage_Costs_2017.pdf. Accessed: Oct. 2019.
[7] M. R. Sarker, D. J. Olsen, and M. A. Ortega-Vazquez, "Co-optimization of distribution transformer aging and energy arbitrage using electric vehicles," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2712–2722, Nov. 2017.
[8] M. A. Ortega-Vazquez, "Optimal scheduling of electric vehicle charging and vehicle-to-grid services at household level including battery degradation and price uncertainty," *IET Gener., Transmiss. Distrib.*, vol. 8, no. 6, pp. 1007–1016, Jun. 2014.
[9] E. A. Grunditz and T. Thiringer, "Performance analysis of current BEVs based on a comprehensive review of specifications," *IEEE Trans. Transp. Electrific.*, vol. 2, no. 3, pp. 270–289, Sep. 2016.
[10] H. Liu, J. Qi, J. Wang, P. Li, C. Li, and H. Wei, "EV dispatch control for supplementary frequency regulation considering the expectation of EV owners," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3763–3772, Jul. 2018.
[11] H. Liu, Z. Hu, Y. Song, and J. Lin, "Decentralized vehicle-to-grid control for primary frequency regulation considering charging demands," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3480–3489, Aug. 2013.
[12] A. Y. S. Lam, K. Leung, and V. O. K. Li, "Capacity estimation for vehicle-to-grid frequency regulation services with smart charging mechanism," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 156–166, Jan. 2016.
[13] S. Izadkhast, P. Garcia-Gonzalez, P. Frìas, L. Ramìrez-Elizondo, and P. Bauer, "An aggregate model of plug-in electric vehicles including distribution network characteristics for primary frequency control," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2987–2998, Jul. 2016.
[14] S. Izadkhast, P. Garcia-Gonzalez, and P. Frías, "An aggregate model of plug-in electric vehicles for primary frequency control," *IEEE Trans. Power Syst.*, vol. 30, no. 3, pp. 1475–1482, May 2015.
[15] K. S. Ko, S. Han, and D. K. Sung, "A new mileage payment for EV aggregators with varying delays in frequency regulation service," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2616–2624, Jul. 2018.
[16] T. Zhang, X. Chen, Z. Yu, X. Zhu, and D. Shi, "A Monte Carlo simulation approach to evaluate service capacities of EV charging and battery swapping stations," *IEEE Trans. Ind. Informat.*, vol. 14, no. 9, pp. 3914–3923, Sep. 2018.
[17] P. Xie, Y. Li, L. Zhu, D. Shi, and X. Duan, "Supplementary automatic generation control using controllable energy storage in electric vehicle battery swapping stations," *IET Gener., Transmiss. Distrib.*, vol. 10, no. 4, pp. 1107–1116, Oct. 2016.
[18] X. Tan, G. Qu, B. Sun, N. Li, and D. H. K. Tsang, "Optimal scheduling of battery charging station serving electric vehicles based on battery swapping," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1372–1384, Mar. 2019.
[19] X. Liu, T. Zhao, S. Yao, C. B. Soh, and P. Wang, "Distributed operation management of battery swapping-charging systems," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5320–5333, Sep. 2019.
[20] PJM Dispatch, "PJM manual 12: Balance operations," 2019. [Online]. Available: https://www.pjm.com/~/media/documents/manuals/m12.ashx. Accessed Mar. 18, 2020.
[21] California Energy Commission, "Electric vehicle charger selection guide," Jan. 2018. [Online]. Available: https://afdc.energy.gov/files/u/publication/EV_Charger_Selection_Guide_2018-01-112.pdf. Accessed: May 27, 2020.
[22] PJM FFRS marketing data. [Online]. Available: http://www.pjm.com/markets-and-operations/data-dictionary.aspx. Accessed: May 27, 2020.
[23] G. He, Q. Chen, C. Kang, P. Pinson, and Q. Xia, "Optimal bidding strategy of battery storage in power markets considering performance-based regulation and battery cycle life," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2359–2367, Sep. 2016.
[24] K. S. Ko, S. Han, and D. K. Sung, "A new mileage payment for EV aggregators with varying delays in frequency regulation service," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2616–2624, Jul. 2018.
[25] K. Kaur, N. Kumar, and M. Singh, "Coordinated power control of electric vehicles for grid frequency support: MILP-based hierarchical control design," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3364–3373, May 2019.
[26] A. Y. S. Lam, K. Leung, and V. O. K. Li, "Capacity estimation for vehicle-to-grid frequency regulation services with smart charging mechanism," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 156–166, Jan. 2016.
[27] M. R. V. Moghadam, R. Zhang, and R. T. B. Ma, "Distributed frequency control via randomized response of electric vehicles in power grid," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 312–324, Jan. 2016.
[28] I. Duggal and B. Venkatesh, "Short-term scheduling of thermal generators and battery storage with depth of discharge-based cost model," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 2110–2118, Jul. 2015.
[29] "Foot Traffic Report for the fuel & convenience retailing industry," GasBuddy and Cuebig, 2018. [Online]. Avaliable: https://www.iab.com/wp-content/uploads/2018/05/GasBuddy-Foot-Traffic-Report-Q1-2018-1.pdf. Accessed: May 27, 2020.
[30] CaliforniaTraffic count data. [Online]. Available: http://pems.dot.ca.gov/. Accessed: May 27, 2020.
[31] X. Wang and G. G. Karady, "Hybrid battery charging strategy for maximizing PV customers' economic benefits," in *Proc. IEEE Power Energy Soc. General Meeting*, Boston, MA, USA, 2016, pp. 1–5.
[32] I. Grondman, L. Busoniu, G. A. D. Lopes, and R. Babuska, "A survey of actor-critic reinforcement learning: Standard and natural policy gradients," *IEEE Trans. Syst., Man, Cybern., Part C (Appl. Rev.)*, vol. 42, no. 6, pp. 1291–1307, Nov. 2012.
[33] R. S. Sutton, D. McAllester, S. Singh, and Y. Mansour, "Policy gradient methods for reinforcement learning with function approximation," in *Proc. 12th Int. Conf. Adv. Neural Inf. Process. Syst.*, Denver, CO, USA, Dec. 1999, pp. 1057–1063.
[34] T. Degris, M. White, and R. S. Sutton, "Off-policy actor-critic," in *Proc. 29th Int. Conf. Mach. Learn.*, Edinburgh, U.K., Jul. 2012, pp. 457–464.
[35] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," 2017, *arXiv:1707.06347*.
[36] J. E. Harlow *et al.*, "A wide range of testing results on an excellent lithium-ion cell chemistry to be used as benchmarks for new battery technologies," *J. Electrochem. Soc.*, vol. 166, no. 13, pp. 3031–3044, Sep. 2019.
[37] E. Redondo-Iglesias, P. Venet, and S. Pelissier, "Efficiency degradation model of lithium-ion batteries for electric vehicles," *IEEE Trans. Ind. Appl.*, vol. 55, no. 2, pp. 1932–1940, Apr. 2019.
[38] S. Benner, "Performance, mileage and the mileage ratio," PJM presentation, Nov. 11, 2015. [Online]. Available: https://www.pjm.com/-/media/committees-groups/task-forces/rmistf/20151111/20151111-item-05-performance-based-regulation-concepts.ashx. Accessed: Jan. 15, 2020.
[39] L. Goldie-Scot, "A behind the scenes take on lithium-ion battery prices," BloombergNEF, Mar. 5, 2019. [Online]. Available: https://about.bnef.com/blog/behind-scenes-take-lithium-ion-battery-prices/. Accessed: Jan. 15, 2020.

Xinan Wang (Student Member, IEEE) received the B.S. degree from Northwestern Polytechnical University, Xi'an, China, in 2013, and the M.S. degree from Arizona State University, Tempe, AZ, USA, in 2016, both in electrical engineering. He is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, Southern Methodist University, Dallas, Texas, USA.

He was a Research Assistant with the AI & System Analytics Group, GEIRI North America, San Jose, CA, USA, in 2016, 2017, and 2019. His current research interests include machine learning applications to power systems, wide-area measurement systems, data analysis, and load modeling.

Jianhui Wang (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from Illinois Institute of Technology, Chicago, Illinois, USA, in 2007. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, Southern Methodist University. He has authored or co-authored more than 300 journal and conference publications, which have been cited for more than 20 000 times by his peers with an H-index of 74.

Dr. Wang is the past Editor-in-Chief of the IEEE TRANSACTIONS ON SMART GRID and an IEEE PES Distinguished Lecturer. He is a Guest Editor of a Proceedings of the IEEE special issue on power grid resilience. He is the recipient of the IEEE PES Power System Operation Committee Prize Paper Award in 2015 and the 2018 Premium Award for Best Paper in IET Cyber–Physical Systems: Theory & Applications. He is a 2018 and 2019 Clarivate Analytics highly cited researcher for production of multiple highly cited papers that rank in the top 1% by citations for field and year in Web of Science. He has been invited to give tutorials and keynote speeches at major conferences including IEEE Conference on Innovative Smart Grid Technologies, IEEE SmartGridComm, IEEE International Conference on Smart Energy Grid Engineering, IEEE International Conference on High Performance and Smart Computing, and 11th International Green Energy Conference.