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/TSG.2017.2666815, IEEE Transactions on Smart Grid

Optimal Charging Operation of Battery Swapping and Charging Stations with QoS Guarantee

Bo Sun, *Student Member, IEEE*, Xiaoqi Tan, *Student Member, IEEE*, and Danny H.K. Tsang, *Fellow, IEEE*

***Abstract*—Motivated by the urgent demand for the electric vehicle (EV) fast refueling technologies, battery swapping and charging stations (BSCSs) are envisioned as a promising solution to provide timely EV refueling services. However, inappropriate battery charging operation in BSCSs can not only incur unnecessary high charging cost but also threaten the reliability of the power grid. In this paper, we aim at obtaining an optimal charging operation policy for a single BSCS to minimize its charging cost while ensuring its quality-of-service. Leveraging the novel queueing network model, we propose to formulate the charging operation problem as a constrained Markov decision process and derive the optimal policy by the standard Lagrangian method and dynamic programming. To avoid the curse of dimensionality in practical large-scale systems, we further analyze the structure of the optimal policy and transform the dynamic programming procedure into an equivalent threshold optimization problem with a discrete separable convex objective function. Numerical results validate our theoretical analysis and the computational efficiency of our proposed algorithms. Our work also shows the impact of the system parameters (e.g., numbers of batteries and chargers) on the average cost under the optimal charging policy, which gives rich insights into the infrastructure planning of future BSCS networks.**

I. Introduction

Electrification of transportation has been considered as a promising solution to mitigate the carbon emission from transportation sector by integrating renewable energy and improve the reliability of the power system by utilizing the energy buffering property of electric vehicles (EVs) [1]. It can be envisioned that EV refueling systems incorporating various technologies (e.g., slow charging, fast charging, battery swapping and wireless charging) will prosper in modern society and serve as the vital joint nodes of transportation and power systems [2]. However, EV refueling systems generally face two crucial challenges: *i*) as service providers, they need to guarantee a certain quality-of-service[1] (QoS) for EV customers to maintain their business profits [3][4]; *ii*) as large electricity consumers, they have to manage their energy consumption profile wisely to reduce their operational cost from the electricity bills and to help maintain the power system reliability [5]. To understand and balance such mutual impact between QoS and operational cost, analytical models that can capture key features of the EV refueling systems are desired. In this paper, we focus on one of the important EV refueling systems, namely, the battery swapping and charging station (BSCS), and study the trade-off between the electricity charging cost and the QoS requirement faced by a single BSCS.

Conceptually a BSCS refuels EVs by swapping their depleted batteries (DBs) with fully-charged batteries (FBs) in store, and charges the DBs to FBs locally with proper charging operation. Agreed with existing works [3]-[6], BSCS is a triple-win solution compared to other EV refueling systems. *First*, EV customers can receive the fastest EV refueling service (e.g., 90 seconds for Tesla battery swapping service [7]) under the pay-for-consumption type business model [3] from BSCSs. Thus, the well-known range anxiety and the high battery purchasing cost can be reduced in the view of EV customers. *Second*, for the power system operators, BSCSs aggregate the uncertain EV charging demand into a single charging entity, whose demand is more predictable and can be reshaped through proper incentive. Therefore, BSCSs are more than large energy consumers, and they can also be considered as high-quality grid-level energy storage. In this regard, the power system reliability can be maintained with less effort. *Third*, superior to other EV refueling systems, the BSCS operators can gain more flexibility from the battery charging operation. The main reason is that BSCSs are able to decouple the EV refueling process into battery swapping procedure and battery charging procedure. Such decoupling gives the BSCSs full control flexibility of the battery charging operation provided enough FBs are available. Thus, the BSCSs can gain profits not just from the payment of EV customers. They can also participate in the wholesale electricity market for buying cheaper and greener energy or providing profitable ancillary services.

Although the concept and technology of BSCSs are already mature, there are still obstacles to the large-scale commercialization of BSCSs. On one hand, BSCSs need a large stock of expensive batteries in the infrastructure planning stage and take the risk for their performance degradation, which leads to much higher capital expense compared to the EV charging stations. On the other hand, due to the lack of standardization [5], batteries of different EV brands vary in physical parameters such as shape, size, and capacity, and may require dedicatedly-designed chargers. Therefore, batteries from different EV brands are not interchangeable. In order to provide services for all types of EVs, a BSCS needs to maintain multiple types of batteries and chargers. Such incompatibility of batteries raises the capital cost and operational difficulties significantly. Thus, it is of great importance to understand the impacts of the numbers of batteries and chargers on the charging operation

This work was supported by the Hong Kong Research Grants Councils General Research Fund under Project 16209814 and Project 16210215.

The authors are with the Department of Electronic and Computer Engineering, the Hong Kong University of Science and Technology, Hong Kong (e-mail: {bsunaa, xtanaa, eetsang}@ust.hk).

[1]QoS refers to the capability of a system to provide better service and can be defined in various ways. Our QoS metric will be specified explicitly later.

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/TSG.2017.2666815, IEEE Transactions on Smart Grid

of BSCSs before planning the infrastructure of the BSCSs.

In this paper, we study the optimal charging operation of the BSCSs and further investigate the impacts of system parameters on BSCSs' operational cost under the optimized charging policies. In summary, the contributions are as follows:

- **An optimal charging operation framework for a single BSCS based on CMDP formulation:** We propose to formulate the charging operation problem as a constrained Markov decision process (CMDP) based on the mixed queueing network model proposed in our prior works [8][12]. Lagrangian method and dynamic programming are utilized to derive the optimal operation policy for this CMDP.
- **Analyzing the optimal policy structure and designing corresponding efficient algorithms:** To avoid the curse of dimensionality from the dynamic programming procedure when solving the CMDP problem, we prove that the optimal policy has a threshold structure under mild assumptions. Leveraging this structural property, we reformulate the dynamic programming procedure as an equivalent threshold vector minimization problem and propose a projected subgradient algorithm to search for the threshold vector efficiently. Our proposed algorithm can reduce the execution time significantly with a small loss of optimality.
- **Investigating the impacts of the numbers of batteries and chargers:** By applying the optimized policies to BSCSs with different system parameters, we numerically evaluate the impacts of the numbers of batteries and chargers on the operational cost and QoS. Our extensive numerical results provide rich insights for further optimization in the infrastructure planning problems on BSCS networks.

II. Related Work

Balancing the operational cost and QoS is the core problem of the BSCS operation. In the literature, the operational cost is typically from the electricity bills [4]-[6], [10]-[14] and/or battery holding cost (e.g., battery degradation cost) [3][6]. The electricity cost can be connected to the charging operation through dynamic electricity price. Specially, this dynamic price may depend on the time [4]-[6], [10]-[11], location [4][9] and/or the instantaneous total charging power of the BSCSs [12][13]-[16]. However, how to appropriately define the QoS metric and quantify its relationship with the charging operation is a nontrivial but inevitable issue deserving further exploration.

A. Existing Approaches and Comparisons

Based on different modeling and formulation techniques, the BSCS charging operation problem mainly falls into two streams in the literature. The first stream of works are static problems in forms of integer/mixed-integer deterministic optimization programs and the QoS metric is defined as the proportion of fulfilled EV swapping requests over the total requests. Early studies [9][10] assume the external information (e.g., swapping request, electricity price, initial SoC of DBs) to be known in advance, which is in fact uncertain. To tackle this uncertainty, [3][6] transform this stochastic information into deterministic parameters with a certain confidence level by leveraging the robust optimization technique. For these works, the probability distributions of the uncertainties are not necessarily required. All the decisions are made at the beginning of the time horizon, only based on the statistics of the system states. Thus, decisions will not change no matter how the uncertainties are realized in the following considered time horizon. In contrast, the second stream of the literature considers sequential decision-making problems, in which decisions are based on the observations of the system state at each decision epoch. Therefore, the solution of this approach is a mapping from the system state to the charging operation, which adapts to the instantaneous realization of the system state. For this stream of works, the QoS metric is defined to be a certain performance metric of the related stochastic models (e.g., queueing/inventory model). In particular, [4] assumes the BSCS has infinite parking spaces for EV customers to wait for swapping services and considers the number of backlogged EVs as the QoS metric. On the other hand, [11][12] consider a finite size of parking spaces and waiting EV customers exceeding the size of the parking spaces are considered to be blocked. Thus, the blocking probability is adopted as the QoS metric.

In a dynamic environment, the sequential decision-making formulation can generally achieve better performance in terms of minimizing the long-term average cost due to its better utilization of the most recent information compared to the static problem formulation. However, the sequential decision-making problems require the probability distributions of the uncertainties, which may sometimes be unavailable and must be based on assumptions. Moreover, unlike static problems, whose solutions can be efficiently computed by leveraging the mature convex or non-convex optimization techniques, sequential decision-making problems are typically computationally difficult and thus hinder practical implementations of large-scale systems.

B. Our Related Works and Comparisons

Our previous works [8][22] adopt a queueing approach to evaluate the performance of the BSCS system without any decision-making process and make contributions to the BSCS infrastructure planning problem. Complementary to [8][22], our existing work [12] and this paper both aim at contributing to the second stream of the charging operation problem of BSCSs.

However, this paper improves [12] in both formulations and solution methods. Specifically, [12] formulates the BSCS charging operation problem as a discrete-time CMDP. The time horizon is discretized into time intervals with equal length and all decisions are made at the beginning of each time interval. This discrete-time model is not able to capture the state transitions within each time interval because the transitions are assumed to occur at the beginning of time intervals promptly. In contrast, this paper formulates the charging operation problem as a continuous-time CMDP, in which decision is made once a system state transition takes place. Thus, our formulation in this paper captures the decision-making process of practical BSCS systems more accurately. In

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/TSG.2017.2666815, IEEE Transactions on Smart Grid

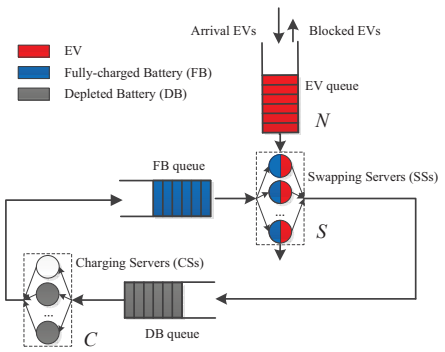

Fig. 1. Queueing network model for an individual BSCS. The EVs form an open queue and the batteries circulate in a closed queue.

terms of solution methods, [12] relaxes the CMDP into an unconstrained MDP by simply adding a weighed blocking probability term to the objective and eliminate the QoS constraint. Then, the unconstrained MDP is solved by standard dynamic programming. However, in this paper, we solve the CMDP directly by Lagrangian method and dynamic programming. Moreover, in order to reduce the computational complexity of the dynamic programming procedure, we analyze the optimal policy structure and design an efficient alternative algorithm to reduce the execution time significantly with a slight loss of optimality.

III. An Overview of the BSCS System Model

We model an individual BSCS as a mixed queueing network, which consists of an open queue of EVs coupled with a closed queue of batteries as shown in Fig. 1.

- **Open queue of EVs**. EVs with battery swapping requests are the customers of the open queue. When the EV queue is not full, EV arrivals will either enter a swapping server (SS), exchange its DB for an FB if available and leave the system or, wait in the parking spaces when there is no idle SS. If the EV queue is fully occupied, EV arrivals will immediately leave for other neighboring BSCSs and these lost customers are considered to be blocked.
- **Closed queue of batteries**. FBs and DBs transit into each other through the SSs or charging servers (CSs). In particular, DBs are charged to become FBs[2] in CSs and then enter the FB queue. FBs are swapped by DBs in SSs and then delivered into the DB queue. Thus, the total number of DBs and FBs in the BSCS keeps constant and hence a fixed number of batteries circulate in a closed queue. Due to the limited number of batteries, both the DB queue and FB queue can be considered to have infinite buffer sizes.

One key feature of this mixed queueing model is that the open queue and the closed queue are coupled together. It can be observed that the EVs (FBs) in the SSs serve as the servers of the FB (EV) queue. Each battery swapping operation in SSs will consume one FB, serve one EV and generate one DB simultaneously. Compared to existing BSCS models [3]-[6], [9]-[11], this mixed queueing model captures *i*) the time for battery swapping operation instead of assuming this operation is finished immediately [3]-[6], [9]-[11]; *ii*) the finite number of parking lots with EV blocking other than considering infinite [3]-[6], [9][10] or zeros parking spaces [11]; and *iii*) the impact of limited numbers of batteries and chargers on the charging operation instead of assuming plenty of chargers (i.e., all DBs are in the chargers) [3][6][10]. These factors are related to the details of charging and swapping operations. Although such factors may be ignored in the long-term infrastructure planning stage, they can affect the real-time operation of BSCSs significantly. Thus, we adopt this mixed queueing model to study the optimal charging operation of BSCSs in this paper. Based on this mixed queueing model, the QoS metric of the BSCS system is defined to be the blocking probability of EV customers.

[2]The FBs refer to the batteries that are full enough for swapping and are not necessarily restricted to be 100% state of charge. The definition of FBs can affect the average charging time of batteries.

IV. CMDP Formulation and Optimal Policy

In this section, we formulate the charging operation problem inside a single BSCS as a stochastic control problem given all the system capacity parameters (i.e., the numbers of parking spaces or EVs allowed in the system N, swapping servers/islands S, chargers C, and batteries B). *The basic trade-off lies in that if too many DBs are charged instantaneously, the charging cost for per unit energy will be raised while charging too few DBs may lead to the shortage of FBs, which will result in higher EV blocking probability.* Thus, the objective of this paper is to find the optimal policy determining how many DBs should be charged concurrently such that the long-term average charging cost is minimized while the EV blocking probability does not exceed a certain value. All the proofs of this paper are presented in the appendix.

A. Assumptions

- **A1: System dynamics.** We first assume that the EV arrivals are Poisson with average rate λ, the time for one swapping operation and the time for charging one battery from DB to FB are both exponentially distributed with average rate ν and μ, respectively. This assumption is common for mathematical tractability of queueing models and has been validated in [8].
- **A2: Homogeneous batteries and chargers.** All the batteries and chargers refer to the same standard. Therefore, only one class of battery (e.g., Tesla battery) is considered and all the chargers have the same constant charging rate r_0 (e.g., 3.3 kW for level-2 EV charging [2]).
- **A3: Uninterruptible charging.** DBs start charging once being put into the chargers and will be charged to FBs without interruptions. This assumption is to avoid the potential deterioration of batteries [19] and simplify the implementation of charging operation in practical BSCS systems in case the smart chargers with control capability are not available.

B. CMDP Formulation

Based on the queueing network model in Fig. 1, the optimal charging operation problem is formulated as a CMDP.

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/TSG.2017.2666815, IEEE Transactions on Smart Grid

1) State space: The state of the BSCS system is defined by a 3-tuple $s = (n, b, c)$, where n and b represent the number of EVs and FBs in the system, respectively, and c indicates the number of busy chargers. Since the total number of batteries inside the BSCS keeps constant, the number of DBs is $B - b$. Then, the state space is denoted by $\mathcal{S} = \{s|b + c = 0, 1, \ldots, B; n = 0, 1, \ldots, N; b = 0, 1, \ldots, B; c = 0, 1, \ldots, C\}$. Note that the non-battery load inside the BSCS, such as lighting, is insignificant compared to the battery-charging load (i.e., $r_0 c$), whereas the total load of a single BSCS represents only a small portion of the total electricity load of the power grid. Thus, we consider the charging operation of the BSCS system is independent of both the non-battery load inside the BSCS and the base load of the power grid.

2) Action space: Decision epochs are chosen as the time instant at which an EV arrives, a battery swapping operation is finished or a battery charging service is completed. Let u be the action determining the number of DBs to be put into the chargers at each decision epoch. Then the action space at state s is denoted by $\mathcal{U}(s) = \{u|u = 0, 1, \ldots, \min\{C - c, B - b - c\}\}$. Note that the action is upper bounded by the number of idle chargers and the number of DBs outside the CSs.

For each state s, its corresponding action is determined by a certain control policy. A policy θ is called stationary and randomized if for any decision epoch, θ maps the current state s to an action $u = \theta(s) \in \mathcal{U}(s)$ with probability $\mathbb{P}^{\theta}(u|s)$. Define the set of stationary randomized policies as Θ. If θ maps state s to only one of the actions in $\mathcal{U}(s)$, θ is called a stationary deterministic policy. The set of the stationary deterministic policies is denoted by Θ_D, and note that $\Theta_D \subseteq \Theta$.

3) Transition kernel: According to **A1**, the sojourn time between the state transition epochs is exponentially distributed. This implies that our problem is a continuous-time CMDP, which can be transformed into an equivalent but easier discrete-time CMDP by uniformization [20]. Let the uniformization rate equal $\gamma = \lambda + C\mu + S\nu$. Then the time interval of the equivalent discrete-time CMDP is $1/\gamma$. Denote the transition rate out of state s after taking action u by $\gamma_{(s,u)} = \lambda + \mu(c + u) + \nu \min\{n, b, S\}$. Then, we have the state transition probability as follows:

$$\mathbb{P}(s'|s,u) = \begin{cases} \frac{\lambda}{\gamma} & s' = (\min\{n+1, N\}, b, c+u), \\ \frac{\mu(c+u)}{\gamma} & s' = (n, b+1, c+u-1), \\ \frac{\nu \min\{n,b,S\}}{\gamma} & s' = (n-1, b-1, c+u), \\ 1 - \frac{\gamma_{(s,u)}}{\gamma} & s' = (n, b, c+u), \\ 0 & \text{otherwise}, \end{cases}$$

where $s, s' \in \mathcal{S}$ denote the state before and after the transition.

4) Cost functions: Electricity charging cost is considered to be the most important operational cost related to the charging operation of BSCSs. By taking the same cost model as [13][14], we assume that the charging cost is quadratic in the total energy consumption of each time interval, which is equivalent to saying that the charging cost within a time interval of $1/\gamma$ is quadratic in the number of busy chargers $c + u$, i.e., $f_c(s, u) = \alpha_0(c + u) + \alpha_1(c + u)^2$. Such charging cost model is also referred to as an approximation of inclining block rate [15], which is a convex piece-wise linear cost function and has been widely used in the literature [12][13]-[16] and in practice [17][18]. More importantly, it follows the fact that the marginal price is strictly increasing with the total demand in the deregulated electricity market. Thus, by using the quadratic cost model, the peak load of large energy consumers, like a BSCS, can be mitigated. Note that based on the aforementioned charging cost function $f_c(s, u)$, the long-term average charging cost under policy θ can be defined as

$$J^{\theta} = \lim_{K\to\infty} \frac{1}{K} \mathbb{E}^{\theta} \left[\sum\nolimits_{k=0}^{K-1} f_c\left(s_k, u_k\right) \right], \tag{1}$$

where k is the index of uniformized time interval.

In order to facilitate the problem solving process, we represent the blocking probability in a long-term average manner as follows:

$$L^{\theta} = \lim_{K\to\infty} \frac{1}{K} \mathbb{E}^{\theta} \left[\sum\nolimits_{k=0}^{K-1} f_l\left(s_k, u_k\right) \right], \tag{2}$$

where $f_l(s, u) = \mathbb{I}_{\{n=N\}}$ is defined to be the EV blocking cost function. $\mathbb{I}_{\{x\}}$ is an indicator function which equals to one if x is true and zero otherwise. Note that the equivalence between equation (2) and the EV blocking probability is due to the well-known PASTA property in queueing theory.

Based on the discussion above, the optimal charging operation problem is formally formulated as

$$(\mathbf{P1}) \quad \min\nolimits_{\theta\in\Theta_{\varepsilon}} J^{\theta} \quad \text{s.t.} \quad L^{\theta} \leq \varepsilon, \tag{3}$$

where ε is a predetermined QoS requirement threshold and $\Theta_{\varepsilon} \subseteq \Theta$ denotes the set of feasible policies. Policy θ is defined to be feasible for **P1** if it can satisfy the QoS constraint $L^{\theta} \leq \varepsilon$. Note that **P1** is an average cost CMDP with finite state and action spaces. Thus, if Θ_{ε} is not empty, there exists a stationary randomized optimal policy θ^*, which minimizes J^{θ} and makes the QoS constraint binding, i.e., $L^{\theta^*} = \varepsilon$ [21].

To further explore the sufficient condition to ensure a nonempty Θ_{ε}, we define a default policy $\theta_d(s) = \min\{C - c, B - b - c\}$, under which the BSCS system puts as many batteries into the chargers as possible. Let J^{θ_d} and L^{θ_d} denote the average charging cost and EV blocking probability respectively under the default policy. Note that the default policy is the most aggressive policy to charge DBs and hence its corresponding blocking probability is the least, i.e., $L^{\theta_d} \leq L^{\theta}, \forall\theta \in \Theta$. Thus, if $\varepsilon < L^{\theta_d}$, Θ_{ε} is empty. The complexity of evaluating L^{θ_d} is equivalent to deriving the steady state probability of the queueing system in Fig. 1. To avoid the trial and error process to understand the system's best QoS L^{θ_d}, we turn to studying the lower bound of L^{θ_d}, which is analytically determined by the system capacity parameters (i.e., N, S, C) and system dynamics (i.e., λ, μ, ν).

Lemma 1. *[22] Given the BSCS system capacity parameters (N, S, C), the lower bound of the EV blocking probability under the default policy, i.e., $\hat{L}^{\theta_d}$, is determined in the following two cases as the number of total batteries approaches infinity:*

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/TSG.2017.2666815, IEEE Transactions on Smart Grid

- ***Case-I:*** *If* $C \leq \lambda(1-\mathbb{P}_{EV}(N,S))/\mu$, $\hat{L}^{\theta_d} = 1 - C\mu/\lambda$. $\mathbb{P}_{EV}(N,S)$ *is the blocking probability of the EV queue when there are infinite FBs and is determined by*

$$\mathbb{P}_{EV}(N,S) = \frac{1}{S^{N-S}S!}\left(\frac{\lambda}{\nu}\right)^N p_0,$$

where $p_0 = \left[\sum_{i=0}^{S}\frac{\lambda^i}{\nu^i i!} + \frac{\lambda^S}{\nu^S S!}\sum_{i=S+1}^{N}\frac{\lambda^{i-S}}{\nu^{i-S}S^{i-S}}\right]^{-1}$.
- ***Case-II:*** *If* $C > \lambda(1-\mathbb{P}_{EV}(N,S))/\mu$, $\hat{L}^{\theta_d} = \mathbb{P}_{EV}(N,S)$.

The detailed proof and physical interpretation of Lemma 1 are presented in our prior work [22] and Appendix A, respectively. Based on Lemma 1, it can be observed that the lower bound has two possibilities depending on the number of chargers. According to [22] and Appendix A, the second case is superior to the first case, and thus should be chosen in the planning of a BSCS. In the rest of the paper, we only consider the BSCS systems that satisfy the condition of Case-II.

C. Optimal Policy of *P1*

In this section, we derive the optimal stationary randomized policy for **P1** by Lagrangian method and dynamic programming. Define the Lagrangian of **P1** as

$$J_\delta^\theta = J^\theta + \delta(L^\theta - \varepsilon) = \lim_{K\to\infty}\frac{1}{K}\mathbb{E}^\theta\Big[\sum_{k=0}^{K-1} f(s_k,u_k;\delta)\Big], \quad (4)$$

where $f(s,u;\delta) = f_c(s,u) + \delta(f_l(s,u)-\varepsilon)$ is the total charging cost per interval and $\delta \geq 0$ is the Lagrangian multiplier. Note that $f(s,u;\delta) = f(s,u)$ if δ is fixed and $f(s,u) = f(n,b,c,u)$. Then, we solve **P1** by the following procedures.

1) Solving the Lagrangian unconstrained MDP problem: For a given δ, we first consider the following problem:

$$(\mathbf{P2}) \quad \min_{\theta\in\Theta_D} \quad J_\delta^\theta. \quad (5)$$

P2 is an unconstrained infinite horizon MDP. Hence an optimal stationary deterministic policy exists for **P2** and can be determined by solving the following Bellman's optimality equation

$$h + V(s) = \min_{u\in\mathcal{U}(s)}\big[f(s,u;\delta) + \sum\nolimits_{s'\in\mathcal{S}}\mathbb{P}(s'|s,u)V(s')\big],$$

where h is uniquely characterized as the corresponding optimal average cost and $V(s)$ is the relative value for state s. If we impose $V(s^0)=0$ for a fixed $s^0\in\mathcal{S}$, V is unique. Based on the standard dynamic programming technique (e.g., relative value iteration (RVI) [23]), the optimal relative value V^* can be iteratively obtained by

$$\begin{aligned} V_i(s) = \min\nolimits_{u\in\mathcal{U}(s)} \quad & \Big[f(s,u;\delta) - V_{i-1}(s^0) + \\ & \sum\nolimits_{s'\in\mathcal{S}}\mathbb{P}(s'|s,u)V_{i-1}(s')\Big], \forall s\in\mathcal{S}, \end{aligned} \quad (6)$$

where i is the iteration index. After convergence, V^* is achieved and the optimal average cost $h^* = V^*(s^0)$. Let θ_δ^* denote the optimal policy of **P2**, then

$$\begin{aligned} \theta_\delta^*(s) = \arg\min\nolimits_{u\in\mathcal{U}(s)} & \big[f(s,u;\delta) - h^* + \\ & \sum\nolimits_{s'\in\mathcal{S}}\mathbb{P}(s'|s,u)V^*(s')\big], \quad \forall s\in\mathcal{S}. \end{aligned} \quad (7)$$

2) Searching for the optimal Lagrangian multipliers: Denote the optimal Lagrangian multiplier and its corresponding optimal policy of **P2** by δ^* and $\theta_{\delta^*}^*$, respectively. Based on the standard results from Lagrangian method in CMDP [21], δ^* is determined by

$$\delta^* = \arg\max\nolimits_{\delta\geq 0}\left(\min\nolimits_{\theta\in\Theta_D} J_\delta^\theta\right) = \arg\max\nolimits_{\delta\geq 0} J_\delta^{\theta_\delta^*}. \quad (8)$$

Moreover, $\theta_{\delta^*}^*$ is also optimal to **P1** if the QoS constraint is strictly binding under policy $\theta_{\delta^*}^*$. This means that solving **P1** is equivalent to searching for δ^* and solving its corresponding **P2**. If there does not exist such a δ^*, which is a common case for CMDP, especially for the discrete state and action spaces, there will exist two optimal Lagrangian multipliers δ^- and δ^+, which satisfy *i)* $L^{\theta_{\delta^-}^*} > \varepsilon$; *ii)* $L^{\theta_{\delta^+}^*} < \varepsilon$; *iii)* $\delta^+ - \delta^- \geq \epsilon, \forall\epsilon > 0$. Specifically, $\theta_{\delta^-}^*$ can achieve lower average charging cost than the optimal policy of **P1** but fails to satisfy the QoS constraint. While for $\theta_{\delta^+}^*$, the QoS constraint is satisfied but it cannot achieve the minimum average charging cost. Then, the optimal policy of **P1** will be a randomized policy which selects $\theta_{\delta^+}^*$ with probability ρ and $\theta_{\delta^-}^*$ with probability $1-\rho$. Specifically, the optimal policy of **P1** can be represented as

$$\theta^* \doteq \rho\theta_{\delta^+}^* + (1-\rho)\theta_{\delta^-}^*. \quad (9)$$

In order to search for the optimal Lagrangian multipliers efficiently, we next show that the Lagrangian multiplier δ and its corresponding optimal policy θ_δ^* satisfy the following monotonicity property.

Lemma 2. ***(Monotonicity Property)*** *The blocking probability* $L^{\theta_\delta^*}$ *is non-increasing in* $\delta\in(0,\infty]$, *and the average charging cost* $J^{\theta_\delta^*}$ *is non-decreasing in* $\delta\in(0,\infty]$.

Based on the monotonicity of $L^{\theta_\delta^*}$ in δ, bisection method is used to search for the optimal Lagrangian multipliers.

3) Obtaining the probability mixture parameter ρ: After obtaining the optimal Lagrangian multipliers δ^+ and δ^- and their corresponding optimal polices $\theta_{\delta^+}^*$ and $\theta_{\delta^-}^*$ by solving **P2**, the remaining part is to derive the probability mixture parameter ρ. Firstly, we can evaluate the blocking probabilities $L^{\theta_{\delta^+}^*}$ and $L^{\theta_{\delta^-}^*}$ under policies $\theta_{\delta^+}^*$ and $\theta_{\delta^-}^*$ by solving the mixed queueing system. In particular, this Markovian system can be solved by linear program (LP) as shown in [8]. Then, ρ is determined by solving the following equation to guarantee the binding of the QoS constraint, namely,

$$L^{\theta^*} = \rho L^{\theta_{\delta^+}^*} + (1-\rho)L^{\theta_{\delta^-}^*} = \varepsilon. \quad (10)$$

The above procedures to obtain the optimal policy θ^* are summarized in Algorithm 1.

V. Policy Structure and Efficient Algorithm

In Algorithm 1, the most time-consuming procedure is solving the unconstrained MDP **P2**. In practical systems (e.g., 50-100 chargers and hundreds of batteries), the dynamic programming based algorithms (e.g., RVI) are unable to obtain the optimal policy numerically due to the curse of dimensionality. To reduce the computational complexity, we explore the policy structure of **P2**, through which an efficient algorithm is proposed to tackle this computational difficulty.

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/TSG.2017.2666815, IEEE Transactions on Smart Grid

Algorithm 1 Computing the optimal policy of **P1**

Input: State space $\mathcal{S}$, action space $\mathcal{U}$, transition kernel $\mathbb{P}(s'|s,u)$, charging cost function $f_c(s,u)$, blocking cost function $f_l(s,u)$ and QoS requirement $\varepsilon \in [L^{\theta_d}, 1)$.
Output: Optimal policy θ^*.
Initialize error tolerance $\epsilon = 10^{-5}$, $\delta^- = 0$, $\delta^+ = 10^4$.
While $\frac{\delta^+ - \delta^-}{2} > \epsilon$ **do**
  $\delta = \frac{\delta^+ + \delta^-}{2}$;
  Determine the optimal policy θ^*_δ by solving **P2**;
  Evaluate the blocking probability $L^{\theta^*_\delta}$;
  if $L^{\theta^*_\delta} = \varepsilon$ **then**
    $\delta^+ = \delta^- = \delta$, $\theta^*_{\delta^+} = \theta^*_{\delta^-} = \theta^*_\delta$.
  else
    if $L^{\theta^*_\delta} < \varepsilon$ **then**
      $\delta^+ = \delta$, $\theta^*_{\delta^+} = \theta^*_\delta$.
    else
      $\delta^- = \delta$, $\theta^*_{\delta^-} = \theta^*_\delta$.
    end if
  end if
end while
Determine ρ by solving equation (10).
$\theta^* \doteq \rho\theta^*_{\delta^+} + (1-\rho)\theta^*_{\delta^-}$.

A. Policy Structure of *P2*

Definition 1. *An order-up-to policy for one dimensional state x is determined by a single threshold R and has the form of*

$$\theta(x) = \begin{cases} R - x & \text{if} \quad x < R \\ 0 & \text{otherwise}. \end{cases} \tag{11}$$

Theorem 1. *The optimal policy of **P2** is an order-up-to policy in c for any fixed (n,b) and can be represented by*

$$\theta^*_\delta(s) = \begin{cases} R^*_{(n,b)} - c & \text{if} \quad c < R^*_{(n,b)} \\ 0 & \text{otherwise}. \end{cases} \tag{12}$$

In particular, the optimal policy parameters $R^*_{(n,b)}$ are determined by the following optimization problem

$$R^*_{(n,b)} = \underset{0 \le \tilde{c} \le \min\{C, B-b\}}{\operatorname{argmin}} f(\tilde{s}, 0; \delta) + \tilde{V}(\tilde{s}), \tag{13}$$

where $\tilde{s} = (n, b, \tilde{c}) = (n, b, c+u)$ is the post-decision state after taking action u and $\tilde{V}(\tilde{s}) = \sum_{s' \in \mathcal{S}} \mathbb{P}(s'|\tilde{s}, 0) V(s')$ denotes the post-decision value function.

Based on Theorem 1, the optimal charging operation policy is order-up-to type and determined by some parameters purely depending on the number of EVs n and the number of FBs b. Therefore, instead of recording the policy as a mapping table (i.e., each state corresponds to a decision), this order-up-to policy simplifies the representation of the optimal policy. Better still, it is much easier for practical implementation. Particularly, at each decision epoch, the optimal decision maintains the total number of busy chargers c up to be a certain value $R^*_{(n,b)}$ for each fixed pair (n,b). This means if $c < R^*_{(n,b)}$ holds, we put $R^*_{(n,b)} - c$ DBs into the chargers. If there are more than $R^*_{(n,b)}$ busy chargers, we stop putting DBs into the chargers. This order-up-to policy structure is illustrated in Fig. 2 (a) in Sec. VI.

Though the optimal parameters $R^*_{(n,b)}$ are the solutions of problem (13), it is non-trivial to solve these optimization problems because the post-decision value functions have no closed-form representations. To derive the computationally efficient algorithm for **P2**, we first have the following corollary.

Corollary 1. ***(Action Space Reduction)*** *For an optimal policy of **P2** with policy parameters $R^*_{(n,b)}$, there exists an alternative optimal policy whose decision is either 0 or 1. In particular, the binary optimal policy is described by*

$$\theta^*_\delta(s) = \begin{cases} 1 & \text{if} \quad c < R^*_{(n,b)} \\ 0 & \text{otherwise}. \end{cases} \tag{14}$$

Corollary 1 has an important computational implication that the searching space for the optimal policy can be significantly reduced especially when the number of chargers is large. Corollary 1 holds because in the steady state of the Markov chain induced by the order-up-to policy, only the states $\left(n, b, R^*_{(n,b)}\right)$, $n = 0, 1, \ldots, N$, and $b = 0, 1, \ldots, B$ have positive stationary probabilities and the other states are transient. Once the system goes away from these steady states, the resulting states are at most one DB away from them. Thus, there exists an alternative binary policy to ensure that the system will eventually enter and circulate in these recurrent states as the order-up-to policy does. Based on Corollary 1, the action space of **P2** is reduced to be $\{0, 1\}$ for the rest of the paper.

Based on our extensive numerical tests on the optimal policies of **P2** with different system parameters, we observe that parameters $R^*_{(n,b)}$ only depend on the value of $n-b$ for almost all the optimal policies. In order to search for the parameters $R^*_{(n,b)}$ efficiently, we make the following assumption.

- **A4:** Policy parameters $R^*_{(n,b)}$ purely depend on $n - b$.

Lemma 3. ***(Threshold Policy)*** *If **A4** holds, the optimal policy of **P2**, $\theta^*_\delta(s)$, is non-decreasing in $n-b$ and is specified by*

$$\theta^*_\delta(s) = \mathbb{I}_{\{n-b > \phi^*_c\}} = \begin{cases} 1 & n - b > \phi^*_c \\ 0 & n - b \le \phi^*_c, \end{cases} \tag{15}$$

where ϕ^*_c is the optimal threshold associated with the number of busy chargers c.

Note that the threshold policy is based on **A4** and is not guaranteed to be an optimal policy of **P2**. However, in order to reduce the computational complexity of **P2**, we propose to restrict the searching policies to threshold policies. As will be shown in Sec. VI, the average cost by applying the optimized threshold polices is rather close to that of the optimal policy (i.e., Fig. 5 (a), Fig. 6 (a) and Fig. 9). Thus, the threshold policy is at least a good approximation of the optimal policy and **A4** is therefore a mild assumption for practical systems. Define a threshold vector $\boldsymbol{\phi} = [\phi_0, \ldots, \phi_C] \in \boldsymbol{\Phi} \cap \mathbb{Z}^{C+1}$, where $\boldsymbol{\Phi} = \{\boldsymbol{\phi} \subseteq [-B, N]^{C+1} | -B \le \phi_0 \le \cdots \le \phi_C \le N\}$. Based on Lemma 3, deriving the optimal policy of **P2** is equivalent to searching for the optimal threshold vector $\boldsymbol{\phi}^*$. Note that $\boldsymbol{\phi}^*$ and $R^*_{(n,b)}$ are both optimal policy parameters,

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/TSG.2017.2666815, IEEE Transactions on Smart Grid

which can specify the optimal policy uniquely and can be transformed into each other.

B. Discrete Convex Reformulation of P2

Definition 2. *A function $f: \mathbb{Z}^M \to \mathbb{R}$ is called discrete separable convex if $f(\boldsymbol{x}) = \sum_{m=1}^{M} f_m(x_m)$, where $f_m: \mathbb{Z} \to \mathbb{R}$ is convex in x_m for all m.*

Separable convexity is a desired property for tractable discrete problems. Next, we show that **P2** can be transformed into an equivalent threshold vector optimization problem with a discrete separable convex objective function. By leveraging this reformulation, a projected subgradient algorithm is used to search for $\boldsymbol{\phi}^*$ efficiently.

Theorem 2. *If Lemma 3 holds, **P2** is equivalent to*

$$\min\nolimits_{\boldsymbol{\phi} \in \boldsymbol{\Phi} \cap \mathbb{Z}^{C+1}} \quad G(\boldsymbol{\phi}) \tag{16}$$

where the objective function

$$G(\boldsymbol{\phi}) = \sum_{s \in \mathcal{S}} \frac{1}{K} \mathbb{E}\Big[\sum_{k=0}^{K-1} f(s_k, \mathbb{I}_{\{n_k - b_k > \phi_{c_k}\}}; \delta) | s_0 = s \Big] \tag{17}$$

is discrete separable convex in $\boldsymbol{\phi}$ with $K \to \infty$.

Due to the discrete separable convexity, the objective function (17) defined on discrete sets can be extended to a continuous convex function $\tilde{G}(\boldsymbol{\psi})$, $\boldsymbol{\psi} \in \boldsymbol{\Phi}$, which has the same minimum and minimizer as $G(\boldsymbol{\phi})$ [24]. Thus, the optimal threshold $\boldsymbol{\phi}^*$ can be achieved by applying subgradient-based algorithms to the extended function $\tilde{G}(\boldsymbol{\psi})$ as shown in Algorithm 2.

Algorithm 2 Solving **P2** by Projected Subgradient Method

Input: Number of iterations I, constant step size β, system capacity parameters N, B, S, C, system dynamics λ, μ, ν.
Output: Threshold vector $\boldsymbol{\phi}^*$.
Initialize $\boldsymbol{\psi}^{(0)} = \mathbf{0}$.
for $i = 1$ **to** I **do**
  Determine the subgradient $\boldsymbol{g}(\boldsymbol{\psi}^{(i-1)})$ by equation (19);
  Step size $\beta^{(i)} = \beta / \left\| \boldsymbol{g}(\boldsymbol{\psi}^{(i-1)}) \right\|_\infty$;
  $\boldsymbol{\psi}^{(i)} = \mathcal{P}_{\Phi}\left(\boldsymbol{\psi}^{(i-1)} - \beta^{(i)} \boldsymbol{g}(\boldsymbol{\psi}^{(i-1)}) \right)$.
end for
Round $\boldsymbol{\psi}^{(I)}$ to its closest integer point $\boldsymbol{\phi}^*$.

The key step in deriving the subgradient of $\tilde{G}(\boldsymbol{\psi})$ at $\boldsymbol{\psi} \in \boldsymbol{\Phi}$ is shown as follows.

1) Constructing $\tilde{G}(\boldsymbol{\psi})$: Let $\boldsymbol{p} = \lfloor \boldsymbol{\psi} \rfloor$ and $\boldsymbol{q} = \boldsymbol{\psi} - \boldsymbol{p}$, where $\lfloor \boldsymbol{x} \rfloor$ denotes the largest integer less than $\boldsymbol{x}$ element by element. Then $\boldsymbol{q} = [q_1, q_2, \ldots, q_{C+1}]^T \in \mathbb{R}^{C+1}$ and $0 \le q_m < 1$, $m = 1, \ldots, C+1$. Define σ as a permutation of the index of $\boldsymbol{q}$ such that $\sigma(m)$ is the index of the m-*th* largest element in $\boldsymbol{q}$. Denote the set of permuted indices by

$$U_0 = \emptyset, \quad U_m = \{\sigma(1), \ldots, \sigma(m)\}, m = 1, \ldots, C+1.$$

Define the characteristic vector of U_m, $m = 0, 1, \ldots, C+1$, as $\chi_{U_m} \in \{0,1\}^{C+1}$, whose j-*th* entry is 1 if $j \in U_m$ and 0 otherwise. Then, $\tilde{G}(\boldsymbol{\psi})$ can be constructed as a linear combination of its surrounding discrete points $G(\boldsymbol{p} + \chi_{U_m})$, $m = 0, 1, \ldots, C+1$. Specifically,

$$\begin{aligned} \tilde{G}(\boldsymbol{\psi}) = (1 - q_{\sigma(1)}) G(\boldsymbol{p}) + q_{\sigma(C+1)} G(\boldsymbol{p} + \chi_{U_{C+1}}) \\ + \sum\nolimits_{m=1}^{C} (q_{\sigma(m)} - q_{\sigma(m+1)}) G(\boldsymbol{p} + \chi_{U_m}). \end{aligned} \tag{18}$$

By this construction, $\tilde{G}(\boldsymbol{\phi}) = G(\boldsymbol{\phi})$, $\forall \boldsymbol{\phi} \in \boldsymbol{\Phi} \cap \mathbb{Z}^{C+1}$ and hence $\tilde{G}(\boldsymbol{\psi})$ can be considered as a piecewise linear extension of $G(\boldsymbol{\phi})$. Furthermore, due to the discrete separable convexity of $G(\boldsymbol{\phi})$, $\tilde{G}(\boldsymbol{\psi})$ is continuously convex [24][25].

2) Evaluating $G(\boldsymbol{\phi})$: Based on equation (18), in order to obtain the subgradient of $\tilde{G}(\boldsymbol{\psi})$ at $\boldsymbol{\psi}$, we need to evaluate $G(\boldsymbol{p} + \chi_{U_m})$ for $m = 0, 1, \ldots, C+1$, namely, the average total cost given the threshold vector $\boldsymbol{p} + \chi_{U_m}$. There are generally two methods to obtain this evaluation. *First*, we can solve an LP [8] or utilize the matrix geometric method [22][26] to achieve the exact value of $G(\boldsymbol{p} + \chi_{U_m})$. By using the exact evaluation, the convergence of the subgradient algorithm is guaranteed due to the convexity of $\tilde{G}(\boldsymbol{\psi})$. *Second*, $G(\boldsymbol{p} + \chi_{U_m})$ can be estimated by simulation if the computational complexity to achieve the exact value is too large [25][27]. In this case, problem (16) can be proved to converge almost surely under standard conditions [25]. In this paper, we use LP to obtain the exact value of $G(\boldsymbol{p} + \chi_{U_m})$. Originally, this LP has $|\mathcal{S}|$ variables and $|\mathcal{S}| + 1$ constraints and its size will increase dramatically with the increase of the size of the state space. Fortunately, the size of such an LP can be further reduced. Particularly, based on the discussion below Corollary 1, only the states $(n, b, R^*_{(n,b)})$ have positive steady state probabilities. Thus, to evaluate the average total cost under a certain policy, the state space of the system can be reduced to two dimensions (n, b). Then, the corresponding LP to evaluate $G(\boldsymbol{p} + \chi_{U_m})$ only has NB variables and $NB + 1$ constraints. In addition, it can be observed that the size of the LP will not increase with the number of chargers, which is important for obtaining the optimal policy efficiently for practical systems with large number of chargers.

3) Deriving the subgradient $\boldsymbol{g}(\boldsymbol{\psi})$: Based on the construction of $\tilde{G}(\boldsymbol{\psi})$ and the evaluation of $G(\boldsymbol{p} + \chi_{U_m})$, the j-*th* entry of the subgradient of $\tilde{G}(\boldsymbol{\psi})$ is determined by

$$\begin{aligned} \boldsymbol{g}_j(\boldsymbol{\psi}) &= \tilde{G}(\boldsymbol{p} + \chi_{U_m}) - \tilde{G}(\boldsymbol{p} + \chi_{U_{m-1}}) \\ &= G(\boldsymbol{p} + \chi_{U_m}) - G(\boldsymbol{p} + \chi_{U_{m-1}}), \end{aligned} \tag{19}$$

where $j = 1, \ldots, C+1$, and $m = \sigma(j)$.

VI. Numerical Results

In this section, we numerically validate our theoretical analysis and algorithms presented in previous sections, and illustrate the impact of the key system parameters (e.g., numbers of batteries and chargers) under optimized charging operations. In our simulations, the fixed system parameters are shown in Table I and others will be specified later.

A. Illustrating the Policy Structure of P2

To illustrate the policy structure of **P2**, we obtain the optimal policy by RVI with the number of batteries $B = 50$,

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/TSG.2017.2666815, IEEE Transactions on Smart Grid

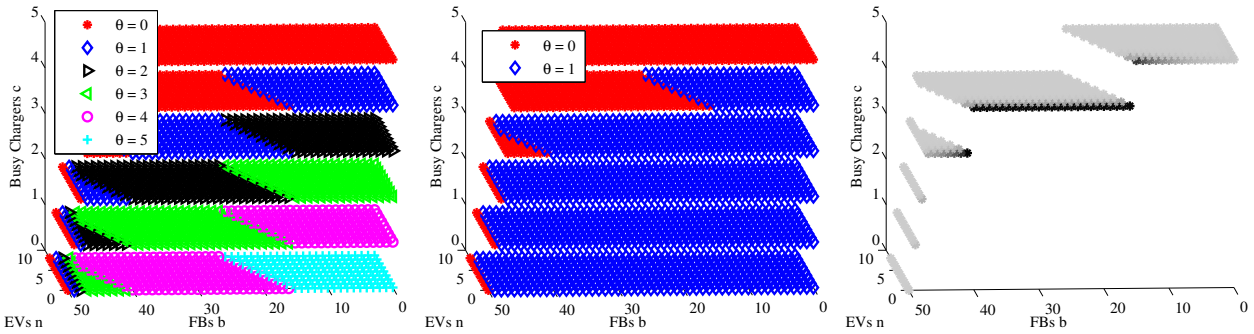

Fig. 2. Illustrating the optimal policy structure of **P2**. (a) Optimal policy structure with full action space. (b) Optimal policy structure after action space reduction. (c) States with positive steady state probabilities under optimal policy. The values of the probabilities are shown by the darkness of the symbol.

TABLE I
SYSTEM PARAMETERS

Number of parking spaces N	8	Arrival rate λ	0.4/min
Number of SSs S	3	Charging rate μ	0.05/min
QoS requirement ε	0.01	Swapping rate ν	1/min

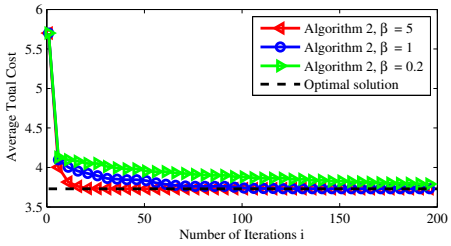

Fig. 3. Convergence performance of average total cost by Algorithm 2.

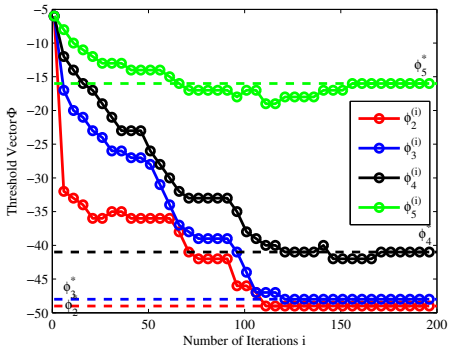

Fig. 4. Convergence performance of optimal threshold $\boldsymbol{\phi}^*$ with $\beta = 1$.

the number of chargers $C = 5$, and Lagrangian multiplier $\delta = 100$. As shown in Fig. 2 (a), different colored symbols represent the optimal decisions for each state. It is clear that the optimal policy is order-up-to type, namely, for each fixed (n, b), $\theta^*(s) = \left[R^*_{(n,b)} - c\right]^+$. Then, we restrict the action space to be $\mathcal{U}(s) = \{0, 1\}, \forall s \in \mathcal{S}$, and obtain the optimal policy as shown in Fig. 2 (b) with the same optimal average charging cost and blocking probability as Fig. 2 (a). This validates the statement in Corollary 1. In addition, it can be observed that for each fixed c, the states with action 0 and 1 are divided by a straight line which purely depends on $n - b$ as stated in Lemma 3. Finally, Fig. 2 (c) shows the states with positive steady state probabilities under the optimal policy. As discussed under Corollary 1, only the recurrent states (i.e., $(n, b, R^*_{(n,b)})$) have positive values. Thus, the average charging cost and blocking probability can be evaluated by solving a linear system only including these states.

B. Illustrating the Performance of Algorithm 2

1) Convergence of Algorithm 2: We start by discussing the method to select the step size parameters in Algorithm 2. Since our algorithm is based on the projected subgradient method, the convergence can be guaranteed by standard choices of step sizes for deterministic problems [28] or stochastic problems [25]. The core idea of the selection is to make sure that the step size is sufficiently small around the optimal value. However, either diminishing step size or small constant step size may lead to slow convergence. In our case, problem (16) is intricately a discrete problem. Therefore, the optimal solution is separate from its surrounding solutions far away compared with the continuous case. Thus, we choose the step size to make sure in each iteration, the solution goes towards the optimal value with a constant step size β in its maximum direction. To avoid the potential oscillation around the optimal solution, Algorithm 2 keeps recording the best solution in each iteration. After applying Algorithm 2 to solve **P2**, Fig. 3 and Fig. 4 illustrate the convergence performance of the average total cost and the threshold vector, respectively. Specifically, in Fig. 3, the optimal value can be achieved within 200 iterations by setting the step size $\beta = 1$. For the case of $\beta = 5$, it converges faster towards the optimal value, but it will finally oscillate in a larger region around the optimal value and the final solution is slightly higher than the optimal value. If $\beta = 0.2$, it converges slowly and will take longer time to reach the optimal value. *In the following simulations, we set* $\beta = 1$.

2) Accuracy and computational complexity of Algorithm 2: We compare the accuracy and complexity between Algorithm 2 and the numerical benchmark RVI method. Accuracy is illustrated by comparing the average total costs obtained from the two algorithms while the complexity is shown

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/TSG.2017.2666815, IEEE Transactions on Smart Grid

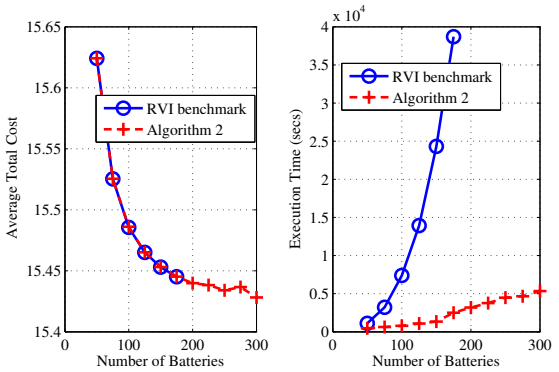

Fig. 5. Comparing the accuracy and complexity between Algorithm 2 and RVI with increase of the number of batteries. $C = 10$.

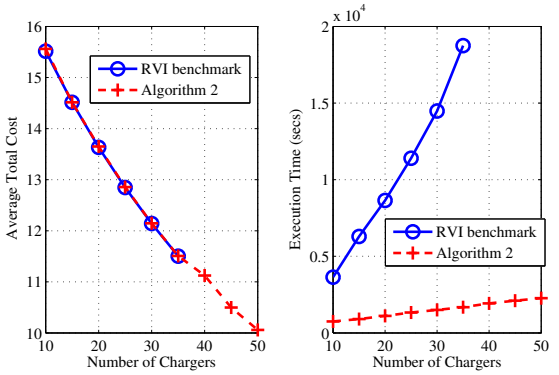

Fig. 6. Comparing the accuracy and complexity between Algorithm 2 and RVI with increase of the number of chargers. $B = 80$.

by comparing the total execution times. Fig. 5 and Fig. 6 illustrate the comparison with varying numbers of batteries and chargers, respectively. With the increase of the system size, the execution time of the RVI method grows exponentially and becomes intractable quickly while Algorithm 2 runs within a reasonable time even for larger systems. Besides, for the tractable cases, the optimal solutions obtained from the two algorithms are relatively close to each other.

3) Impact of Lagrangian Multiplier δ: We illustrate the impact of the Lagrangian multiplier on the average charging cost $J^{\theta_\delta^*}$, blocking probability $L^{\theta_\delta^*}$, and Lagrangian polynomial $J_\delta^{\theta_\delta^*}$ when applying the Lagrangian method. As defined in equation (4), Lagrangian polynomial is a weighed sum of the average charging cost and blocking probability. In particular, the Lagrangian multiplier is the weight of the blocking probability and transforms the blocking probability into a monetary value to make it comparable with the average charging cost. Thus, δ can be interpreted as the marginal cost of blocking probability. By this interpretation, the Lagrangian polynomial represents the average total cost (i.e., charging cost plus EV blocking cost) of the BSCS. In Figs. 7 (a) and (b), the average charging cost is non-decreasing in δ and the blocking probability is non-increasing in δ, because the increase of δ will raise the importance of the EV blocking, and hence results in more aggressive policies with higher cost to charge batteries so that the blocking probability can be reduced. In addition, these two figures verify the monotonicity property of the Lagrangian multiplier, which is proved theoretically in Lemma 2. Finally, Fig. 7 (c) shows that the Lagrangian polynomial is maximized at the optimal Lagrangian multiplier. For different settings of the QoS requirement ε, the optimal Lagrangian multiplier is different, in order to guarantee the QoS requirement. Specifically, stringent QoS requirement (i.e., small ε) requires high marginal cost of blocking probability (i.e., large δ).

C. Illustrating the Optimal Policy of *P1*

In this section, we investigate the impact of the numbers of batteries and chargers on the average charging cost by applying the optimized policies from Algorithm 1. Specially, we solve **P2** by RVI and Algorithm 2, and the resulting policies are respectively named as *optimal policy* and *Alg2-based policy*.

1) Trade-off between average charging cost and blocking probability: Fig. 8 shows how the blocking probability affects the average charging cost for different numbers of chargers and batteries. Because the QoS constraint will be strictly binding under the optimal policy in **P1**, i.e., $L^{\theta^*} = \varepsilon$, we increase the QoS requirement ε gradually and derive its corresponding optimal average charging cost by solving **P1**. For all the settings, average charging cost decreases with the increase of blocking probability. Moreover, increasing the numbers of chargers and/or batteries can achieve a lower curve in the figure, which means that the BSCSs with more batteries or chargers can achieve the same blocking probability with lower average charging cost. More importantly, compared to batteries, increasing the number of chargers can significantly reduce the average charging cost. For example, all the three curves with circle markers have the same number of batteries (i.e., $B = 80$). These curves have large drop on the average charging cost with the increase of the number of chargers. However, comparing the curves with circle and triangle markers, we observe that for the curves with the same number of chargers, increasing the number of batteries can only slightly reduce the charging cost if the QoS requirement is stringent (i.e., blocking probability is small). Otherwise, the number of batteries can hardly make any improvement in terms of cost reduction. In fact, the top two curves, which only have 8 chargers, are even not able to guarantee very small blocking probability due to the lack of chargers.

2) Impact of the numbers of chargers and batteries: As shown in Fig. 9, the average charging cost decreases with the increase of the numbers of batteries and chargers. Moreover, the number of chargers has a larger influence on the average charging cost. This phenomenon indicates that when designing the capacity parameters of BSCSs and taking the future operational cost into consideration, the system planners need to pay more attention to the number of chargers instead of the total number of batteries. In addition, the performance of the Alg2-based policy is close to the optimal benchmark. Thus, for large systems, Alg2-based policy is a good alternative to the optimal policy in terms of accuracy and complexity.

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/TSG.2017.2666815, IEEE Transactions on Smart Grid

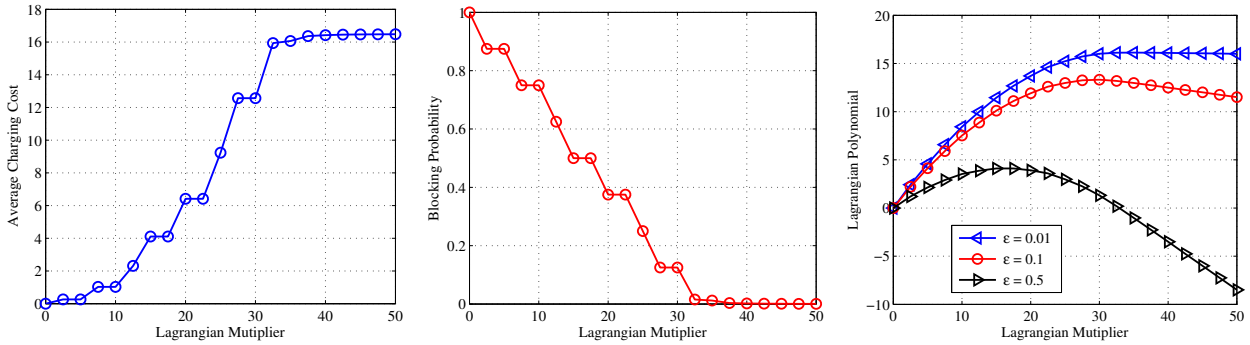

Fig. 7. Illustrating the impact of Lagrangian multiplier on (a) average charging cost; (b) blocking probability; and (c) Lagrangian polynomial (average total cost). The numbers of chargers and batteries are set to be $C = 10$ and $B = 80$.

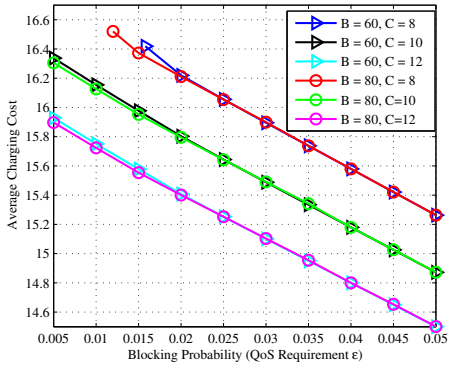

Fig. 8. Trade-off between average charging cost and blocking probability under optimal policy.

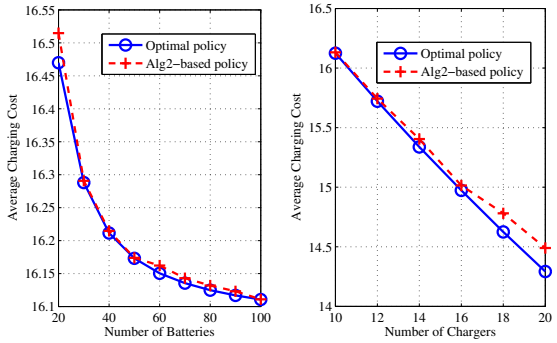

Fig. 9. Illustrating the average charging cost under the optimal policy and Alg2-based policy. The EV blocking probability is guaranteed to be 0.01. (a) $C = 10$. (b) $B = 80$.

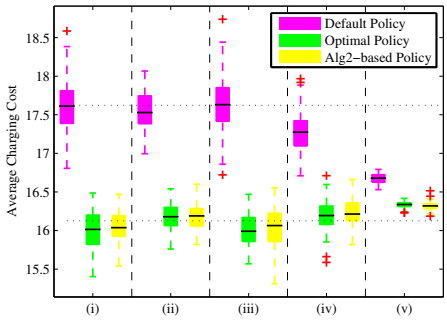

Fig. 10. Boxplot statistics of the average charging cost evaluated by Monte Carlo simulation. The optimal policy and Alg2-based policy perform better than default policies for all cases. The average charging costs under the optimized polices are close when only one of the three distributions deviates from exponential distribution.

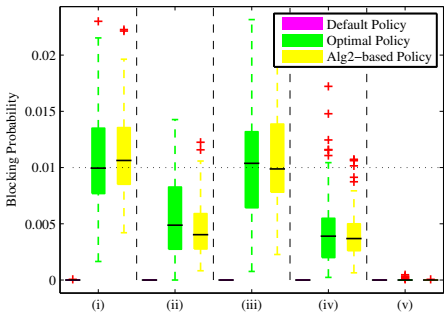

Fig. 11. Boxplot statistics of the blocking probability evaluated by Monte Carlo simulation. Default polices achieve lower blocking probabilities than the optimal and Alg2-based polices in all cases. The blocking probabilities under optimized polices are around or below the target QoS requirement 0.01 for all the cases.

D. Identifying the Effects of System Dynamics

The theoretical analysis and proposed algorithms are mainly based on **A1**, which assumes that all the dynamics of the BSCS system are exponentially distributed. Once the distributions in practice deviate from this assumption, our proposed algorithms are not necessarily able to guarantee the same performance as designed. However, in order to investigate how our model will be affected by the distributions, we apply the policies derived based on **A1** to the systems with

TABLE II
DISTRIBUTIONS OF SYSTEM DYNAMICS

Case	EV interarrival time	swapping time	charging time
(i)	Expon(2.5)	Expon(1)	Expon(20)
(ii)	Expon(2.5)	Expon(1)	Uniform(15,25)
(iii)	Expon(2.5)	Constant(1)	Expon(20)
(iv)	Uniform(1.5,3.5)	Expon(1)	Expon(20)
(v)	Uniform(1.5,3.5)	Constant(1)	Uniform(15,25)

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/TSG.2017.2666815, IEEE Transactions on Smart Grid

different distributions of EV arrival, times for battery swapping and charging operations, and evaluate the average charging cost and blocking probability by Monte Carlo simulation. In particular, the distribution combinations of the five testing cases are shown in Table[3] II. In the simulation, the numbers of the chargers and batteries are set to be $C = 10$ and $B = 80$, respectively. The blocking probability requirement is set to be 0.01. For each case, the simulations are performed for a time horizon of 30 days under *default policy*, *optimal policy* and *Alg2-based policy*, respectively, and are repeated 100 times. The simulation results are shown in Figs. 10 and 11. Case (*i*) is the benchmark case, which adopts the distributions consistent with **A1**. Therefore, the performance of the optimized polices in this case match the theoretic results, namely, guaranteeing the blocking probability to be 0.01 and minimizing the average charging cost. We maintain the means of different distributions in the following cases to be the same as those in case (*i*). Due to the lack of historical data for the charging time distribution in existing BSCSs, we choose to test the uniform distribution (i.e., case (*ii*)), which has been adopted in [6], to make a comparison with the exponential distribution in case (*i*). It can be observed that in case (*ii*), the blocking probability is lower than the QoS requirement by sacrificing the charging cost performance. In practice, the swapping operation may be finished by swapping robots with a constant time. Therefore, the constant-time swapping operation is tested in case (*iii*). By comparing cases (*i*) and (*iii*), the average charging cost and blocking probability are rather close. For the EV interarrival time distribution, existing works [4][11][12] with sequential decision-making formulations accept the Poisson arrival assumption. For comparison, we choose the uniform distributed interarrival time in case (*iv*). Similar to the performance of case (*ii*), case (*iv*) achieves lower blocking probability with higher average charging cost. This indicates that in both cases, there exists better policies that can achieve a lower average charging cost and restrict the blocking probability to be 0.01. Finally, if all the distributions deviate from **A1**, the performance degrades significantly as shown in case (*v*).

VII. Conclusions

In this paper, we have formulated the QoS-guaranteed optimal charging operation problem for BSCSs as a CMDP based on a novel queueing network model. To tackle the computational difficulty of solving the CMDP, we have carefully examined the optimal policy structure and designed a corresponding efficient algorithm. The accuracy and complexity of our proposed algorithm have been compared with the standard RVI methods and its superior performance has been validated by extensive numerical evaluations. Finally, our numerical results have shown that the number of chargers in the BSCS has a larger impact in reducing the average charging cost when the system is operated under QoS-guaranteed optimal policies.

[3]In the table, all cases are in unit of minutes. Expon(x) refers to exponential distribution with average rate x, Constant(x) represents deterministic distribution of constant x and Uniform(x,y) stands for uniform distribution in $[x, y]$.

Appendix A
Interpretation of Lemma 1

Basically, Lemma 1 states that as the number of total batteries approaches infinity, the EV blocking probability will converge asymptoticly to a lower bound $\hat{L}^{\theta_d}$. Moreover, depending on the relationship of the *maximum battery charging service rate* $C\mu$ and the *maximum EV arrival rate* $\lambda(1 - \mathbb{P}_{EV}(N, S))$, $\hat{L}^{\theta_d}$ appears in two different forms. In Case-I, $C\mu$ is less than or equal to $\lambda(1 - \mathbb{P}_{EV}(N, S))$, and hence the battery charging process is the bottleneck of BSCS's QoS. Therefore, there must be a positive probability that FBs are not enough to offer the battery swapping service, which leads to EVs queueing up and severe blocking events. In contrast, Case-II represents the system with plenty of chargers so that $C\mu$ is larger than $\lambda(1-\mathbb{P}_{EV}(N, S))$. In this case, $\hat{L}^{\theta_d}$ is determined by the battery swapping process and is equal to the blocking probability of the $M/M/S/N$ queueing system (EV queue). This means that in Case-II, it is with probability 1 that there exist enough FBs for the battery swapping service. Thus, from the EVs' point of view, the BSCS is just an $M/M/S/N$ queue. More importantly, $\hat{L}^{\theta_d}$ in Case-I is always larger than that in Case-II. This fact implies that if the system parameters satisfy the conditions of Case-II, BSCS can achieve a smaller blocking probability when B is large enough and hence has more flexibility to control the battery charging process. Thus, the settings of the system parameters in Case-II are more appealing in practice and will be the focus of our paper.

Appendix B
Proof of Lemma 2

For any $\delta > 0$ and $\xi > 0$, we have

$$J_{\delta+\xi}^{\theta_\delta^*} - J_\delta^{\theta_\delta^*} \geq J_{\delta+\xi}^{\theta_{\delta+\xi}^*} - J_\delta^{\theta_\delta^*} \geq J_{\delta+\xi}^{\theta_{\delta+\xi}^*} - J_\delta^{\theta_{\delta+\xi}^*}, \tag{20}$$

where the two inequalities are due to the optimality of $J_\delta^{\theta_\delta^*}$ for **P2**. Since $J_{\delta+\xi}^{\theta_\delta^*} - J_\delta^{\theta_\delta^*} = \xi(L^{\theta_\delta^*} - \varepsilon)$, inequality (20) can be converted to

$$\xi(L^{\theta_\delta^*} - \varepsilon) \geq J_{\delta+\xi}^{\theta_{\delta+\xi}^*} - J_\delta^{\theta_\delta^*} \geq \xi(L^{\theta_{\delta+\xi}^*} - \varepsilon). \tag{21}$$

Thus, $L^{\theta_\delta^*} \geq L^{\theta_{\delta+\xi}^*}$, namely, $L^{\theta_\delta^*}$ is non-increasing in δ.

Again, due to the optimality of $J_\delta^{\theta_\delta^*}$, we have

$$\begin{aligned} J_\delta^{\theta_{\delta+\xi}^*} - J_\delta^{\theta_\delta^*} &= J^{\theta_{\delta+\xi}^*} + \delta(L^{\theta_{\delta+\xi}^*} - \varepsilon) - \left[J^{\theta_\delta^*} + \delta(L^{\theta_\delta^*} - \varepsilon)\right] \\ &= J^{\theta_{\delta+\xi}^*} - J^{\theta_\delta^*} + \delta(L^{\theta_{\delta+\xi}^*} - L^{\theta_\delta^*}) \geq 0. \end{aligned}$$

Since $\delta(L^{\theta_{\delta+\xi}^*} - L^{\theta_\delta^*}) \leq 0$, then $J^{\theta_{\delta+\xi}^*} \geq J^{\theta_\delta^*}$, $\forall \delta > 0$.

Appendix C
Proof of Theorem 1

For the convenience of representation, we omit the parameter δ in the equations within this proof. For each state $s \in \mathcal{S}$ and action $u \in \mathcal{U}(s)$, we denote its post-decision value function in iteration i as

$$\tilde{V}_i(\tilde{s}) = \sum\nolimits_{s' \in \mathcal{S}} \mathbb{P}(s'|\tilde{s}, 0) V_{i-1}(s'). \tag{22}$$

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/TSG.2017.2666815, IEEE Transactions on Smart Grid

In addition, define an ancillary function

$$Q_i(s,u) = f(\tilde{s},0) + \tilde{V}_i(\tilde{s}) = f(s,u) + \tilde{V}_i(n,b,c+u), \quad (23)$$

where $f(\tilde{s},0) = f(s,u)$ since $\tilde{s} = (n,b,c+u)$. Then, we have

$$V_i(s) = \min_{u\in\mathcal{U}(s)} Q_i(s,u). \quad (24)$$

Proposition 1. *The following statements hold for each iteration i: (i) $\tilde{V}_i(\tilde{s})$ is convex in $\tilde{c}$; (ii) $Q_i(s,u)$ is convex in (c,u); (iii) $V_i(s)$ is convex in c.*

Proof. We first prove (*i*) of the above proposition by induction. To show the convexity of $\tilde{V}_i(\tilde{s})$ in $\tilde{c}$, we equivalently prove $\tilde{V}_i(n,b,\tilde{c}+1) - \tilde{V}_i(n,b,\tilde{c})$ is non-decreasing in $\tilde{c}$. Due to $V_0(s) = 0$, $\forall s \in \mathcal{S}$, it is clear that $\tilde{V}_1(\tilde{s}) = 0$, $\forall \tilde{s} \in \mathcal{S}$. Thus, $\tilde{V}_1(n,b,\tilde{c}+1) - \tilde{V}_1(n,b,\tilde{c})$ is non-decreasing in $\tilde{c}$. Assume $\tilde{V}_{i-1}(n,b,\tilde{c}+1) - \tilde{V}_{i-1}(n,b,\tilde{c})$ is non-decreasing in $\tilde{c}$. Then after substituting $\mathbb{P}(s'|\tilde{s},0)$ according to the transition kernel defined in Sec. IV-B into equation (22) and manipulating some terms, we have

$$\begin{aligned}
\tilde{V}_i(n,b,\tilde{c}+1) - \tilde{V}_i(n,b,\tilde{c}) = \\
&\frac{\lambda}{\gamma}[V_{i-1}(\min\{n+1,N\},b,\tilde{c}+1) - \\
&\qquad V_{i-1}(\min\{n+1,N\},b,\tilde{c})] \\
&+ \frac{\mu\tilde{c}}{\gamma}\left[V_{i-1}(n,b+1,\tilde{c}) - V_{i-1}(n,b+1,\tilde{c}-1)\right] \\
&+ \frac{\nu\min\{n,b,S\}}{\gamma}[V_{i-1}(n-1,b-1,\tilde{c}+1) - \\
&\qquad V_{i-1}(n-1,b-1,\tilde{c})] \\
&+ (1 - \frac{\gamma_{(n,b,\tilde{c}+1,0)}}{\gamma})[V_{i-1}(n,b,\tilde{c}+1) - V_{i-1}(n,b,\tilde{c})] \\
&+ \frac{\mu}{\gamma}[V_{i-1}(n,b+1,\tilde{c}) - V_{i-1}(n,b,\tilde{c})].
\end{aligned}$$

To prove the above equation is non-decreasing in $\tilde{c}$, we need to show that $V_{i-1}(n,b,c)$ is convex in c and the last term $V_{i-1}(n,b+1,\tilde{c}) - V_{i-1}(n,b,\tilde{c})$ is non-decreasing in $\tilde{c}$.

Firstly, by the induction hypothesis, $\tilde{V}_{i-1}(\tilde{s})$ is convex in $\tilde{c}$. Based on equation (23), we can easily verify $Q_{i-1}(s,u)$ is convex in (c,u). Due to the fact that convexity is preserved under minimization, $V_{i-1}(s) = \min_{u\in\mathcal{U}(s)} Q_{i-1}(s,u)$ is convex in c. Secondly, note that $V_{i-1}(n,b,\tilde{c}) - V_{i-1}(n,b+1,\tilde{c})$ represents the total expected cost with one less FB in initial states. This cost difference stems from the future EV blocking cost due to the shortage of FBs. In physical systems, such expected blocking cost decreases with the increase of busy servers c because more busy servers means faster FB supply and hence the blocking cost difference induced by one less FB will be reduced. This validates that $V_{i-1}(n,b,\tilde{c}) - V_{i-1}(n,b+1,\tilde{c})$ is non-increasing in $\tilde{c}$ and thus the reverse is non-decreasing in $\tilde{c}$. By summarizing the above two points, $\tilde{V}_i(n,b,\tilde{c}+1) - \tilde{V}_i(n,b,\tilde{c})$ is non-decreasing in $\tilde{c}$.

Finally, (*ii*) and (*iii*) can be easily verified based on (*i*). □

After substituting $\tilde{c} = c + u$ and $u = \tilde{c} - c$ into equation (24), the optimal policy can be described as

$$\theta^*(s) = \left(\operatorname*{argmin}_{c\le\tilde{c}\le\min\{C,B-b\}} f(n,b,\tilde{c},0) + \tilde{V}(n,b,\tilde{c})\right) - c. \quad (25)$$

Then we show that $\theta^*(s)$ can be uniquely determined by the optimal solution of the following relaxed problem

$$R^*_{(n,b)} = \operatorname*{argmin}_{0\le\tilde{c}\le\min\{C,B-b\}} f(n,b,\tilde{c},0) + \tilde{V}(n,b,\tilde{c}). \quad (26)$$

Based on Proposition 1, $f(n,b,\tilde{c},0) + \tilde{V}(n,b,\tilde{c})$ is convex in $\tilde{c}$ for $0 \le \tilde{c} \le \min\{C, B-b\}$. Then, we can have *i*) if $c \le R^*_{(n,b)} \le \min\{C, B-b\}$, then the optimal policy is $\theta^*(s) = R^*_{(n,b)} - c$; *ii*) if $R^*_{(n,b)} < c$, due to the convexity, $f(n,b,\tilde{c},0) + \tilde{V}(n,b,\tilde{c})$ is non-decreasing in $\tilde{c}$ for $c \le \tilde{c} \le \min\{C, B-b\}$, then the optimal decision is $\theta^*(s) = c - c = 0$. Thus, the optimal policy of **P2** is in the form of order-up-to type policy.

APPENDIX D
PROOF OF LEMMA 3

Define $\mathcal{S}_d = \{(n,b,c) | n - b = d, d = -B, \ldots, N\}$ as the set of states with the same $n-b$. Based on **A4**, let $Q(d,c,u) = Q(s,u)$, $\forall s \in \mathcal{S}_d$, where $Q(s,u)$ is defined in equation (23). Then, in order to prove $\theta^*(s)$ is non-decreasing in $n-b$, we just need to show $Q(d,c,u)$ is submodular in (d,u), $\forall c$ [23].

Definition 3. ***(Submodularity)*** *A function $f(x,y)$ is submodular in (x,y) if for $x^+ \ge x^-$ and $y^+ \ge y^-$,*

$$f(x^+,y^+) - f(x^+,y^-) \le f(x^-,y^+) - f(x^-,y^-). \quad (27)$$

Referring to the definition of submodularity, we need to prove $Q(d,c,u+1) - Q(d,c,u)$ is monotonically non-increasing in d. In particular,

$$\begin{aligned}
Q(d,c,u+1) - Q(d,c,u) = \\
&f(n,b,c,u+1) - f(n,b,c,u) \\
&+ \tilde{V}(n,b,c+u+1) - \tilde{V}(n,b,c+u).
\end{aligned} \quad (28)$$

Since $f(n,b,c,u+1) - f(n,b,c,u)$ is irrelevant to $n-b$, we just need to show $\tilde{V}(n,b,c+u+1) - \tilde{V}(n,b,c+u)$ is non-increasing in $n-b$. The proof of this argument is by induction, whose procedure is similar to the proof of Proposition 1. The only difference is that we need to show $V_{i-1}(n,b+1,c) - V_{i-1}(n,b,c)$ is non-decreasing in b. As discussed in Proposition 1, $V_{i-1}(n,b,c) - V_{i-1}(n,b+1,c)$ represents the potential blocking cost due to one less FB in store. Such expected future cost is non-increasing in b because more FB in stock indicates a smaller future blocking probability. Then, $V_{i-1}(n,b+1,c) - V_{i-1}(n,b,c)$ is non-decreasing in b and hence $Q(d,c,u+1) - Q(d,c,u)$ is non-increasing in d. Thus, $Q(d,c,u)$ is submodular in (d,u).

APPENDIX E
PROOF OF THEOREM 2

It is straightforward that **P2** is equivalent to the following optimization problem

$$\min_{\theta} \quad \sum\nolimits_{s\in\mathcal{S}} J^{\theta}_{\delta}(s), \quad (29)$$

where $J^{\theta}_{\delta}(s)$ is defined in equation (4) with initial state $s \in \mathcal{S}$. Based on Lemma 3, problems (16) and (29) are equivalent. Thus, problem (16) is a reformulation of **P2**. In order to prove the discrete convexity of the objective function (17)

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/TSG.2017.2666815, IEEE Transactions on Smart Grid

in $\boldsymbol{\phi}$, we first note that $G(\boldsymbol{\phi}) = \sum_{s\in\mathcal{S}} Q(s, \mathbb{I}_{\{n-b>\phi_c\}}) = (N+1)\sum_{c=0}^{C}\sum_{d=-B}^{N} Q(d,c,\mathbb{I}_{\{d>\phi_c\}})$. In addition, we define $\hat{Q}(c,\phi_c) = \sum_{d=-B}^{N} Q(d,c,\mathbb{I}_{\{d>\phi_c\}})$. Then, we have

$$\begin{aligned}\hat{Q}(c,\phi_c+1)+\hat{Q}(c,\phi_c-1)-2\hat{Q}(c,\phi_c)\\ =\Big[Q(\phi_c-1,c,1)-Q(\phi_c-1,c,0)\Big]\\ -\Big[Q(\phi_c,c,1)-Q(\phi_c,c,0)\Big]\geq 0,\end{aligned} \tag{30}$$

where the inequality is due to the submodularity of $Q(d,c,u)$ in (d,u) shown in the proof of Lemma 3. Thus, $\hat{Q}(c,\phi_c)$ is convex in ϕ_c. In addition, $G(\boldsymbol{\phi}) = (N+1)\sum_c \hat{Q}(c,\phi_c)$. Thus, $G(\boldsymbol{\phi})$ is discrete separable convex in $\boldsymbol{\phi}$.

REFERENCES

[1] M. D. Galus, M. G. Vay, T. Krause and G. Andersson, "The role of electric vehicles in smart grids," *Wiley Interdiscipl. Rev.: Energy Environ.*, vol. 2, pp. 384-400, 2013.

[2] M. Yilmaz and P. T. Krein, "Review of battery charger topologies, charging power levels, and infrastructure for plug-in electric and hybrid vehicles," *IEEE Trans. Power Electron.*, vol. 28, pp. 2151-2169, 2012.

[3] H. Mak, Y. Rong and Z. M. Shen, "Infrastructure planning for electric vehicles with battery swapping," *Manage. Sci.*, vol. 59, pp. 1557-1575, 2013.

[4] F. Schneider, U. Thonemann, and D. Klahjan, "Optimization of battery charging and purchasing at electric vehicle battery swap stations," submitted to *Transp. Sci.*, 2015.

[5] Y. Zheng , Z. Y. Dong , Y. Xu , K. Meng , J. H. Zhao and J. Qiu,, "Electric vehicle battery charging/swap stations in distribution systems: comparison study and optimal planning," *IEEE Trans. Power Syst.*, vol. 29, no. 1, pp. 221-229, 2014.

[6] M. Sarker, H. Pandzic, and M. Ortega-Vazquez, "Optimal operation and services scheduling for an electric vehicle BSS," *IEEE Trans. Power Syst.*, vol. 30, no. 2, pp. 901-910, 2015.

[7] K. Bullis, "Why Tesla thinks it can make battery swapping work," MIT Technology Review, June, 2013, [Online]. Available: http://www.technologyreview.com/news/516276/ why-tesla-thinks-it-can-make-battery-swapping-work/.

[8] X. Tan, B. Sun, and D.H.K. Tsang, "Queueing network models for EV charging station with battery swapping," *in Proc. of IEEE SmartGridComm 2014*, Venice, Nov. 2014.

[9] S. Nurre, R. Bent, F. Pan, and T. Sharkey, "Managing operations of plugin hybrid electric vehicle (PHEV) exchange stations for use with a smart grid," *Energy Policy*, vol. 67, pp. 364-377, April 2014.

[10] T. Raviv, "The battery switching station scheduling problem," *Oper. Res. Letters*, vol. 40, no. 6, pp. 546-550, June 2012.

[11] Q. Dong, D. Niyato, P.Wang, and Z. Han, "The PHEV charging scheduling and power supply optimization for charging stations," *IEEE Trans. Veh. Technol.*, vol. 65, no. 2, pp. 566-580, Feb. 2016

[12] B. Sun, X. Tan and D.H.K Tsang, "Optimal operation of BSSs with QoS guarantee," *in Proc. of IEEE SmartGridComm 2014*, Venice, Nov. 2014.

[13] Y. He, B. Venkatesh and L. Guan "Optimal scheduling for charging and discharging of electric vehicles," *IEEE Trans. Smart Grid*, vol. 3, pp. 1095-1105, 2012.

[14] Z. Ma, D. Callaway, I. Hiskens, "Decentralized charging control for large populations of plug-in electric vehicles: Application of the Nash certainty equivalence principle", *in Proc. IEEE Int. Conf. Control Appl.*, 2010.

[15] A. H. Mohsenian-Rad and A. Leon-Garcia, "Optimal residential load control with price prediction in real-time electricity pricing environments," *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 120-133, 2010.

[16] W. Tang, S. Bi and Y. Zhang, "Online coordinated charging decision algorithm for electric vehicles without future information," *IEEE Trans. Smart Grid*, vol. 5, no. 6, pp. 2810-2824, 2014.

[17] HK Electric, *Block Rate Tariff*, 2016, [Online]. Available: https://www.hkelectric.com/en/customer-services/billing-payment-electricity-tariffs/residential-tariff

[18] BC Hydro, *Electricity Rates*, 2016, [Online]. Available: https://www.bchydro.com/accounts-billing/rates-energy-use/electricity-rates/business-rates.html.

[19] L. Gan, U. Topcu and S. H. Low, "Stochastic distributed protocol for electric vehicle charging with discrete charging rate," in *Proc. IEEE Power Eng. Soc. Gen. Meet.*, San Diego, USA, Jul., 2012.

[20] D. P. Bertsekas, *Dynamic Programming and Optimal Control, vol. II*, Athena Scientific, 1995.

[21] E. Altman, *Constrained Markov Decision Processes*. Boca Raton, FL: Chapman and Hall/CRC Press, 1999.

[22] X. Tan, B. Sun, Y. Wu and D.H.K. Tsang, "Asymptotic performance evaluation of a battery swapping and charging station for electric vehicles," submitted to *Performance Evaluation*, 2016.

[23] W. B. Powell, *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, John Wiley & Sons, 2007.

[24] K. Murota, "Note on multimodularity and L-convexity," *Math. Oper. Res.*, vol. 30, no. 3, pp. 658-661, 2005.

[25] E. Lim, 'Stochastic approximation over multidimensional discrete sets with applications to inventory systems and admission control of queueing networks," *ACM Trans. Modeling and Comput. Simulation*, vol. 22, no. 4, pp. 19:1-19:23, Nov. 2012.

[26] M. F. Neuts, *Matrix-geometric Solutions in Stochastic Models: An Algorithmic Approach*, Dover Publications, 1994.

[27] N. Ding, P. Sadeghi and R.A. Kennedy, "Discrete convexity and stochastic approximation for cross-layer onoff transmission control," *IEEE Trans. Wireless Commun.*, vol. 15, pp. 389-400, 2016.

[28] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 1995.

Bo SUN (S'14) received the B.E. degree in electronic and information engineering from Harbin Institute of Technology, Harbin, China, in 2013. He is currently pursing the Ph.D. degree in electronic and computer engineering at the Hong Kong University of Science and Technology. His research interests are on queueing theory, Markov decision process and discrete convex analysis with applications in energy systems.

Xiaoqi Tan (S'12) is currently a Ph.D. student in the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology. He received his B.E. degree from the Department of Information and Telecommunication Engineering (first class honor), Xi'an Jiaotong University, Xi'an, China, 2012. He is interested in developing analytic techniques and efficient algorithms in stochastic modelling, queueing theory, optimization and control, with current research focusing on applying these models and techniques to the fields of smart grids and power systems.

Danny H.K. Tsang (M'82-SM'00-F'12) received the Ph.D. degree in electrical engineering from the University of Pennsylvania, Philadelphia, in 1989. Since 1992, he has been with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong, where he is currently a Professor. His current research interests include Internet quality of service, peer-to-peer (P2P) video streaming, cloud computing, cognitive radio networks, and smart grids. Dr. Tsang served as a Guest Editor for the IEEE Journal of Selected Areas in Communications Special Issue on Advances in P2P Streaming Systems, an Associate Editor for the Journal of Optical Networking published by the Optical Society of America, and a Guest Editor of the IEEE Systems Journal. He currently serves as Technical Editor for the IEEE Communications Magazine.