Stochastic Distributed Protocol for Electric Vehicle Charging with Discrete Charging Rate

Lingwen Gan, Ufuk Topcu, *Member, IEEE,* and Steven H. Low, *Fellow, IEEE*

Abstract—To address the grid-side challenges associated with the anticipated high electric vehicle (EV) penetration level, various charging protocols have been proposed in the literature. Most if not all of these protocols assume continuous charging rates and allow intermittent charging. However, due to charging technology limitations, EVs can only be charged at a fixed rate, and the intermittency in charging shortens the battery lifespan. We consider these charging requirements, and formulate EV charging scheduling as a discrete optimization problem.

We propose a stochastic distributed algorithm to approximately solve the optimal EV charging scheduling problem in an iterative procedure. In each iteration, the transformer receives charging profiles computed by the EVs in the previous iteration, and broadcasts the corresponding normalized total demand to the EVs; each EV generates a probability distribution over its potential charging profiles accordingly, and samples from the distribution to obtain a new charging profile. We prove that this stochastic algorithm *almost surely* converges to one of its *equilibrium* charging profiles, and each of its equilibrium charging profiles has a *negligible* sub-optimality ratio. Case studies corroborate our theoretical results.

Index Terms—distributed control, discrete optimization, stochastic algorithm, electric vehicle charging.

Notation

n	EV index, $n = 1, \dots, N$
t	time slot index, $t = 1, \dots, T$
D	non-EV demand profile
r_n	charging profile of EV n
r	charging profile of all EVs
$\mathcal{F}_n$	the set of potential charging profiles for EV n
x'	the transpose of a matrix (or vector) x
$\langle x, y \rangle$	$\langle x, y \rangle := x'y$
$\|x\|$	$\|x\| := \sqrt{\langle x, x \rangle}$

I. Introduction

ELECTRIC vehicles (EVs) offer significant potential for increasing energy efficiency in transportation, reducing greenhouse gas emissions, and relieving reliance on foreign oil [1]. Currently, several types of EVs are either already in the U.S. market or about to enter [2], and electrification of transportation is at the forefront of many research and development agendas [3]. On the other hand, the potential comes with a multitude of challenges including those in the integration into the electric power grid. For example, EV charging potentially amplifies peak electricity load [4], increases power losses and voltage variations [6], and reduces the distribution circuit and transformer lifespan [5].

The authors are with Engineering and Applied Sciences at the California Institute of Technology, e-mail: lgan@caltech.edu; utopcu@cds.caltech.edu; slow@caltech.edu.

Many studies demonstrate that adopting "smart" charging strategies can mitigate some of the integration challenges, defer infrastructure investment needed otherwise, and even stabilize the grid. For example, scheduling EV charging so that the aggregated EV load fills the overnight electricity load valley may reduce daily cycling of the power plants [11]. Furthermore, the energy stored in the EVs may be utilized as an ancillary service resource [7] for regulating voltage profiles, ride-through support for fault protection, and even compensating fluctuating renewable energy generation [8].

A potential EV charging control architecture is shown in Figure 1 [9]. In this architecture, there is no communication between the transformers and the substation, so that the control is within systems composed of one transformer and several EVs. For each such system, computation and communication devices will be installed at the transformer and at the charger of each EV to enable distributed control. The transformer will broadcast signals like the price of electricity usage, to guide the EVs in deciding when to charge their batteries. Centralized control at the transformer will not be implemented due to its high computational complexity.
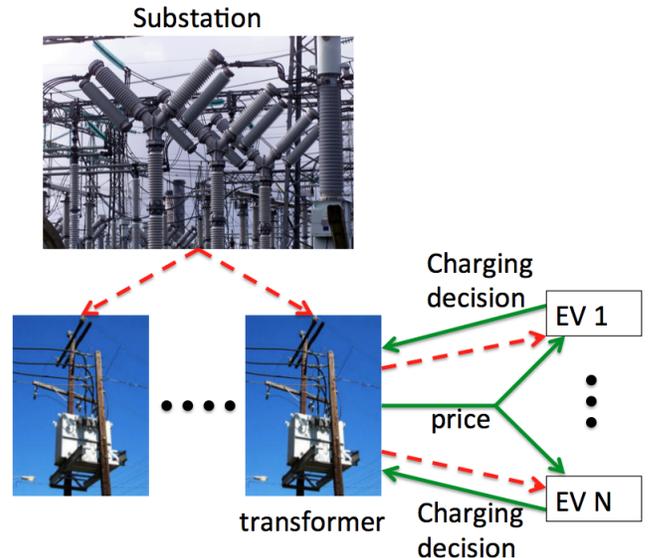

Fig. 1. A potential EV charging control architecture. Dashed red lines represent power flow, and solid green lines represent communication.

Studies on EV charging scheduling roughly fall into two categories: centralized charging control [6], [8], [10], and decentralized charging control [11], [12]. Reference [6], [8], [10] study centralized control strategies that minimize power losses, load variance, or maximize load factor. Reference [11]

proposes a decentralized charging algorithm to flatten the total demand with performance guarantee in the *homogeneous case*, where all the EVs are essentially the same. Reference [12] proposes a decentralized charging algorithm to flatten the total demand with performance guarantee in all cases. All the aforementioned works make the assumptions that (1) EV charging can be intermittent; and (2) the charging rate takes values in a continuous range. In reality, an EV either does not charge or charges at a given rate, and interruptions in charging should be avoided to protect the battery.

The contributions of the current paper include three folds. First, we take into account the facts that charging rate is fixed and charging should not be interrupted, and formulate EV charging scheduling as a discrete optimization problem, where the objective is to flatten the total demand. Second, we compare deterministic and stochastic algorithms, and conclude by establishing a counterexample that any deterministic algorithm cannot in general flatten the total demand under the new charging constraints. Third, motivated by this fact, we propose a stochastic distributed algorithm to approximately solve the optimal EV charging scheduling problem. We prove that our algorithm *almost surely* converges to one of its *equilibrium* charging profiles, and each of its equilibrium charging profiles has a *negligible* sub-optimality ratio.

The rest of the paper is organized as follows. Section II formulates EV charging scheduling as a discrete optimization problem. Section III proposes a distributed stochastic solution algorithm and provides performance guarantees for it. Case studies are presented in section IV, and conclusions are summarized in section V. For clarity, we skip the proofs in the main sections, and present them in the appendix.

II. PROBLEM FORMULATION

Consider the scenario where a transformer negotiates with N electric vehicles (EVs) on their charging profiles for a day-ahead scheduling. Partition the scheduling horizon into T time slots of equal duration, for example, 15 minutes. Due to the limitations in EV charging technology, EVs can only be charged at a fixed rate. For example, household charging rate is 3.3kW [13]. Besides, intermittency in charging shortens the lifespan of a battery (which is very expansive). Hence, we should not interrupt charging an EV in the day-ahead scheduling. Consequently, we make the following assumptions.

A1 An EV consumes a fixed power when it is charging.
A2 EV charging cannot be interrupted.

Then, an EV can only choose the time it starts charging, after which it has to be charged at a fixed rate until its battery is full.

A3 An EV starts charging at the beginning of a time slot.

We can relieve the restriction imposed by assumption A3 by choosing a finer partition of the scheduling horizon. We make assumption A3 so that an EV has only finite optional times to start charging.

Remark 1: A few sample, hypothetical charging profiles satisfying assumptions A1 and A2 are illustrated in Figure 2(a). In practice, more realistic charging profiles look like the ones in Figure 2(b): charging rate ramps up when an EV starts
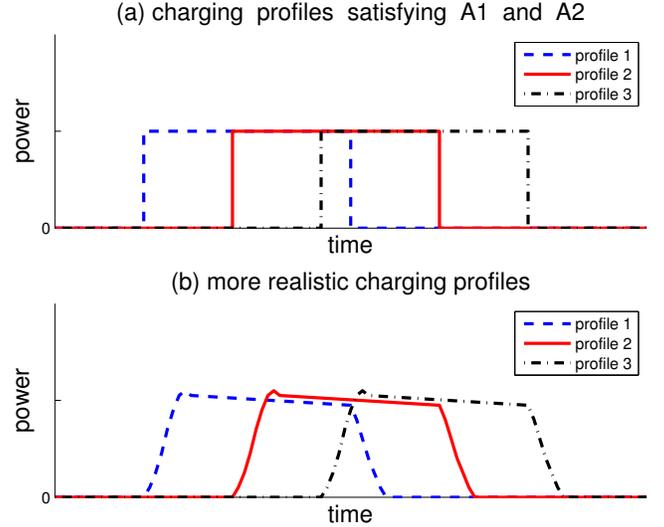

Fig. 2. Hypothetical charging profiles. In this paper, we consider the charging profiles in (a), but extension to the charging profiles in (b) is straight forward.

charging, slowly decreases as the battery gets full, and ramps down when the charging process is about to end. In this paper, we consider the charging profiles in Figure 2(a), but extension to the charging profiles in Figure 2(b) is straight forward.

Let $t = 1, \ldots, T$ denote the time slots and $n = 1, \ldots, N$ denote the EVs. Assume that the charging rate of EV n in time slot t is a constant, and denote it by $r_n(t)$. Define $r_n := (r_n(1), \ldots, r_n(T))'$ as the charging profile of EV n, and $r := \{r_1, \ldots, r_N\}$ as the charging profile of all EVs.

As an example, suppose that EV n becomes available for charging at (the beginning of) time slot 4, and needs to finish charging by (the beginning of) time slot 20. Its battery requires 6 time-slot durations to get fully charged (at its constant charging rate). We call time slot 4 the *plug-in time*, time slot 20 the *deadline*, and the number 6 of time-slot durations *charging length*. We denote the plug-in time, deadline, and charging length of EV n by $plug_n$, $dead_n$, and $length_n$ respectively. Define $S_n := (plug_n, dead_n, length_n)$, and call it the *specification* of EV n.

In this example, the potential time slots to start charging are slots $4, 5, \ldots, 14$. More generally, the set $\mathcal{T}_n$ of potential starting time slots for EV n is

$$\mathcal{T}_n := \{t_n \in \mathbb{Z} \mid plug_n \leq t_n \leq dead_n - length_n\}.$$

Each potential starting time slot $t_n \in \mathcal{T}_n$ corresponds to a potential charging profile, and we index these potential charging profiles by $a = 1, \ldots, A_n$, where

$$A_n := dead_n - length_n - plug_n + 1$$

is the total number of potential starting time slots. Let $r_{n,a}$ denote the ath potential charging profile for EV n, and

$$\mathcal{F}_n := \{r_{n,a} \mid a = 1, \ldots, A_n\}$$

denote the set of all potential charging profiles for EV n.

We consider the case where the transformer knows the inelastic non-EV demand, and aims to flatten the total demand through scheduling the time slots EVs start charging. Let $D(t)$ denote the non-EV demand at time slot t and $D := (D(1), \ldots, D(T))'$ denote the non-EV demand profile. In practice, D is well predicted at the transformer level, and charging control aims to avoid overheating the transformer. Since we shape the total demand $D + \sum_n r_n$ only through shifting EVs' energy consumption, the total energy consumption $\sum_t (D(t) + \sum_n r_n(t))$ is fixed. Consequently, flattening the total demand helps cooling the transformer. According to [12], flattening the total demand $D + \sum_n r_n$ is equivalent to minimizing the objective function

$$L(r) = L(r_1, \ldots, r_N) := \left\| D + \sum_n r_n \right\|^2. \qquad (1)$$

This leads to the following optimization problem, which we call **optimal discrete charging** (ODC) hereafter.

$$\mathbf{ODC} \begin{cases} \underset{r_1, \ldots, r_N}{\text{minimize}} & L(r_1, \ldots, r_N) \\ \text{subject to} & r_n \in \mathcal{F}_n, \ n = 1, \ldots, N. \end{cases}$$

Each decision variable r_n takes discrete values in $\mathcal{F}_n$, making ODC hard to solve [16]. Furthermore, we need a distributed algorithm to avoid high-complexity centralized computation, adding extra difficulty.

Remark 2: In this paper, we consider the objective function in (1) for clarity, but the results extend easily to any objective function of the form

$$\tilde{L}(r_1, \ldots, r_N) := \sum_t C_t \left(\sum_n r_n(t) \right),$$

where for each t, C_t is a convex function with the second derivative $C_t''(\sum_n r_n(t)) \leq \beta_t < \infty$ for all $r_n(t)$. Note that

- if $C_t(\sum_n r_n(t)) := (D(t) + \sum_n r_n(t))^2$, the function $\tilde{L}$ coincides with (1);
- if we want the total demand to track a given profile $G(t)$, then we can define C_t as

$$C_t \left(\sum_n r_n(t) \right) := \left(D(t) + \sum_n r_n(t) - G(t) \right)^2,$$

and use the corresponding $\tilde{L}$ as our objective function.

III. Algorithm

In this section, we propose a distributed solution algorithm for the discrete optimization problem ODC.

A. Information Flow Pattern

We focus on distributed algorithms in which the transformer and the EVs exchange information in multiple iterations to agree on a charging profile r. As mentioned in the introduction, the transformer broadcasts control signals (like the price of electricity usage), and the EVs choose slots to start charging accordingly. The resulting information flow pattern is depicted in Figure 3. The transformer knows the non-EV demand D and each EV n knows its own specification

$S_n = (plug_n, dead_n, length_n)$. For each variable, let the superscript k denote its respective value in iteration k. For example, r_n^k denotes the charging profile r_n computed by EV n in iteration k. In iteration $k = 1, 2, \ldots$:

F1 The transformer receives charging profiles $r_1^{k-1}, \ldots, r_N^{k-1}$ computed by the EVs in the previous iteration[1], and calculates a control signal d^k based on these charging profiles and the non-EV demand D according to

$$d^k = f\left(D, r_1^{k-1}, \ldots, r_N^{k-1}\right), \qquad (2)$$

where the map f is to be designed. After calculating d^k, the transformer broadcasts d^k to all the EVs.

F2 Each EV n receives the broadcast signal d^k computed by the transformer in the current iteration, and calculates a new charging profile r_n^k based on d^k, its specification S_n, and its previously calculated charging profiles r_n^{k-1} according to

$$r_n^k = g\left(d^k, S_n, r_n^{k-1}\right), \qquad (3)$$

where the map g is to be designed. After calculating r_n^k, EV n reports r_n^k to the transformer.
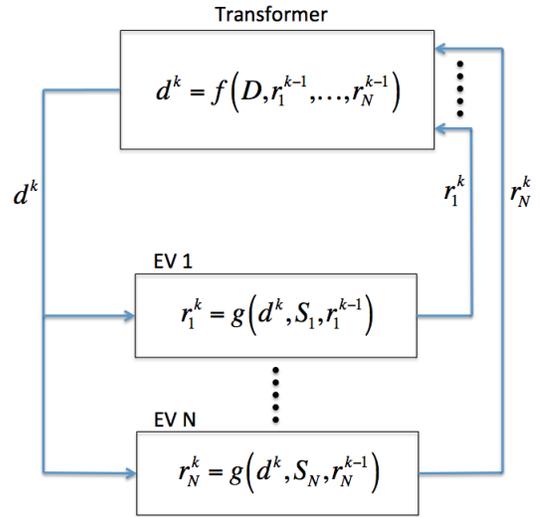

Fig. 3. Information flow pattern in the proposed iterative, distributed decision-making process. The transformer is considered to know the non-EV demand D and each EV n knows its specification $S_n = (plug_n, dead_n, length_n)$.

Remark 3: The transformer uses the broadcast signal d^k to guide the EVs in choosing their charging profiles, so that objectives like flattening the total demand can be achieved. An example of the signal d^k can be the price of electricity usage.

B. Deterministic v.s. Stochastic Algorithms

With the information flow pattern in section III-A, deterministic algorithms will not in general be able to flatten the total demand. We start by formalizing the notions of deterministic and stochastic algorithms.

Definition 1: A map $f : \mathcal{X} \to \mathcal{Y}$ is deterministic provided that for all $x \in \mathcal{X}$, $f(x) = y$ for some $y \in \mathcal{Y}$.

---

[1]If iteration k is the first iteration EV n computes r_n, then $r_n^{k-1} := 0$.

Definition 2: *Let $\mathcal{Y} = \{y_1, \ldots, y_A\}$ be a finite set. The vector $p := (p_1, \ldots, p_A)'$ is a probability distribution over $\mathcal{Y}$ with p_a being the mass probability of y_a for $a = 1, \ldots, A$, provided that (i) $p_a \geq 0$ for $a = 1, \ldots, A$; (ii) $\sum_{a=1}^{A} p_a = 1$.*

Definition 3: *Let $\mathcal{Y}$ be a finite set, $f : \mathcal{X} \to \mathcal{Y}$ be a map, and $\Theta_{\mathcal{Y}}$ be the set of all probability distributions over $\mathcal{Y}$. The map f is stochastic provided that for all $x \in \mathcal{X}$, $f(x)$ is a random variable with some probability distribution $\theta(x) \in \Theta_{\mathcal{Y}}$, that is parametrized by x.*

Given an input $x \in \mathcal{X}$, a deterministic map f^{det} gives a deterministic output $y \in \mathcal{Y}$, and a stochastic map f^{sto} gives an output that is a random variable, which takes values in $\mathcal{Y}$.

Definition 4: *A distributed algorithm using the information flow pattern described in F1 and F2 is deterministic (stochastic), provided that the map g in (3) is deterministic (stochastic).*

In order to prove the claim at the beginning of section III-B, consider the *homogeneous* case, where the specifications S_n of all the EVs are the same.

Proposition 1: *In the homogeneous case, every deterministic algorithm using the information flow pattern described in F1 and F2 will get the same charging profile for all EVs in all iterations, i.e., $r_m^k = r_n^k$ holds for any iteration $k \geq 1$ and any two EVs m and n.*

Proof: See Appendix A. □

Proposition 1 implies that in general, deterministic algorithms will not be able to spread out the time slots EVs start charging. Consequently, the total demand cannot be flattened. However, in the well-studied *continuous charging rate setting* where the charging rate of an EV can take continuous values, various deterministic algorithms have been proposed to flatten the total demand with optimality guarantees in the homogeneous case [11], [12]. Where does this "contradiction" comes from? We call our setting where an EV can only be charged at a fixed rate the *discrete charging rate setting*.

In the continuous charging rate setting in the homogeneous case, though deterministic algorithms get all the EVs to charge at the same time with the same rate, they can change the total demand at slot t through adjusting EVs' charging rates at time t. However, in the discrete charging rate setting, the only way to change the total demand at slot t is to change the number of EVs that charge at slot t. For deterministic algorithms, this number is either 0 or N. Consequently, deterministic algorithms cannot flatten the total demand. In conclusion, there is no contradiction between Proposition 1 and the aforementioned results on using deterministic algorithms to flatten the total demand.

Proposition 1 only discusses deterministic algorithms using the information flow pattern described by F1 and F2. It is possible to find a deterministic algorithm that flattens the total demand using other information flow patterns. For instance, consider the following information flow pattern: in each iteration k, the transformer uses the most recently calculated $r_1, \ldots, r_N$ to compute the broadcast signal d^k; and only one of the EVs (in turn) computes a new charging profile, and reports it to the transformer. In the homogeneous case, a deterministic algorithm that uses this information flow pattern and flattens the total demand can be found. However, such an information flow pattern requires higher communication overhead and is

likely to converge slower.

Besides the disadvantage of deterministic algorithms in Proposition 1, there is an advantage of stochastic algorithms. We illustrate the advantage through an example where there are a large number of homogeneous EVs. For each EV n, define a charging profile matrix

$$R_n := (r_{n,1}, \ldots, r_{n,A_n}) \in \mathbb{R}^{T \times A_n},$$

whose columns are all potential charging profiles for EV n. Define Θ_n as the set of all probability distributions over $\mathcal{F}_n = \{r_{n,1}, \ldots, r_{n,A_n}\}$.

When the EVs are homogeneous, we neglect the subscript n in $r_{n,a}$, A_n, R_n, $\mathcal{F}_n$, Θ_n, and denote them by r_a, A, R, $\mathcal{F}$, Θ without ambiguity. Suppose that all the EVs choose their charging profiles according to the same probability distribution $p = (p_1, \ldots, p_A)'$ over $\mathcal{F}$, where p_a is the mass probability of r_a for $a = 1, \ldots, A$. By the law of large numbers, the average charging profile of the EVs is approximately $\sum_a p_a r_a = Rp$. Note that Rp is a convex combination of the charging profiles in $\mathcal{F}$, and as p traverses Θ, Rp traverses the convex hull $conv(\mathcal{F})$ of $\mathcal{F}$. Hence, we can approximately relax the constraint $r_n \in \mathcal{F}_n$ to $r_n \in conv(\mathcal{F}_n)$ by exploiting stochasticity, and then problem ODC is turned into a convex optimization problem [15].

Due to the aforementioned disadvantage of deterministic algorithms and advantage of stochastic algorithms, we develop a stochastic algorithm in the following section.

C. A Stochastic Algorithm

In this section, we propose a distributed stochastic algorithm that *almost surely* converges to one of its *equilibrium* charging profiles, and show that each of its equilibrium charging profiles has a *negligible* sub-optimality ratio. Recalling the definition of R_n and Θ_n in section III-B, we give the

Distributed Stochastic Charging (DSC) Algorithm:
The transformer knows the non-EV demand D and the number N of EVs; each EV n knows its set $\mathcal{F}_n$ of potential charging profiles. Initialize $r_n^0 := 0$ for all n.
Transformer's algorithm: At iteration $k = 1, 2, \ldots$,

1) the transformer receives charging profiles $r_1^{k-1}, \ldots, r_N^{k-1}$ calculated by the EVs in the previous iteration[2], and computes the corresponding normalized total demand

$$d^k := \frac{D + \sum_n r_n^{k-1}}{N};$$

2) the transformer broadcasts d^k to all EVs.

EV n's algorithm $(n = 1, \ldots, N)$: At iteration $k = 1, 2, \ldots$,

1) EV n receives broadcast signal d^k from the transformer, and computes probability distribution

$$p_n^k := \underset{p \in \Theta_n}{\operatorname{argmin}} \left\| \frac{N}{N-1}(d^k - r_n^{k-1}) + R_n p \right\|^2; \quad (4)$$

2) EV n draws a sample r_n^k according to the probability distribution p_n^k, and reports r_n^k to the transformer.

---

[2]If $k = 1$, use the initialized value $r_1^0 = 0, \ldots, r_N^0 = 0$.

After a fixed number K (for example, $K = 20$) of iterations, terminate the iterative procedure. The transformer and the EVs agree on charging profile $r^K = \{r_1^K, \ldots, r_N^K\}$.

Remark 4: Note that EV n needs to know N in calculating (4). In practice, the transformer can broadcast N together with d^k in iteration 1. Besides, since the term $\frac{N}{N-1}$ is close to 1 when N is large, the EVs may substitute $\frac{N}{N-1}$ by 1 in calculating (4) so that they do not need to know N.

After giving the distributed stochastic algorithm DSC, we formalize the statements at the beginning of section III-C. We start with the definition of an equilibrium charging profile.

Definition 5: *For a given stochastic algorithm, let r^k denote its charging profile in iteration k. A charging profile r^e is an equilibrium for the given stochastic algorithm, provided that the escape probability $P^{escape}(r^e)$ at r^e is zero:*

$$P^{escape}(r^e) := \sup_{k \geq 0} \Pr\{r^{k+1} \neq r^k \mid r^k = r^e\} = 0. \quad (5)$$

If the transition probability $\Pr\{r^{k+1} \mid r^k\}$ is k-independent, then (5) is equivalent to

$$P^{escape}(r^e) := \Pr\{r^1 \neq r^0 \mid r^0 = r^e\} = 0.$$

In whichever iteration k, if r^k equals an equilibrium r^e, then $r^{k+1}, r^{k+2}, \ldots$ will stay at r^e with probability 1.

Definition 6: *A stochastic algorithm is convergent, if*

(a) *it has equilibrium charging profiles;*

(b) *the sequence $r^1, r^2, \ldots$ of charging profiles generated by the stochastic algorithm converges to one of its equilibrium charging profiles with probability 1. This type of convergence is also called almost surely convergence.*

Theorem 1: *Algorithm DSC has equilibrium charging profiles.*

Proof: See Appendix D. $\square$

All the proofs for this paper can be found in [**?**]. Note that prominent stochastic algorithms like simulated annealing [18] and genetic algorithm [19] do not have equilibrium points r^e, since for both algorithms, the following probability is nonzero at any iteration k and any point r:

$$\Pr\{r^{k+1} \neq r^k \mid r^k = r\} > 0. \quad (6)$$

One of the reasons these algorithms satisfy (6) is that they aim to escape from non-optimal points, but it turns out that they escape from optimal points as well. The methodology in these stochastic algorithms is that as iterations evolve, the mass probability of the optimal point gets larger. Eventually, the optimal point (if it is unique) will have a mass probability close to 1. This process usually takes a large number of iterations.

Instead of taking a large number of iterations to find an optimal charging profile for the discrete optimization problem ODC (which is NP hard), Algorithm DSC turns to finding a good enough sub-optimal charging profile through a small number of iterations: it sets the escape probability to be 0 at some good enough (quantified later) sub-optimal charging profiles. At the expense of small sub-optimality, Algorithm DSC obtains existence of equilibrium charging profiles, which is essential for a stochastic algorithm to be convergent. Furthermore, the following theorem shows that Algorithm DSC is indeed convergent.

Theorem 2: *Algorithm DSC is convergent, i.e., the sequence $r^1, r^2, \ldots$ generated by Algorithm DSC converges to (one of its) equilibrium charging profiles with probability 1.*

Proof: See Appendix E. $\square$

After studying the convergence of Algorithm DSC, we evaluate the sub-optimality of its equilibrium charging profiles.

Theorem 3: *Let $r^e = \{r_1^e, \ldots, r_N^e\}$ be an arbitrary equilibrium charging profile of Algorithm DSC, then r^e is a Nash equilibrium. That is, if only one EV n changes its charging profile from r_n^e to $\tilde{r}_n$, then the cost function does not decrease, i.e.,*

$$\left\| D + \sum_m r_m^e \right\| \leq \left\| D + \sum_{m \neq n} r_m^e + \tilde{r}_n \right\|.$$

Proof: See Appendix F. $\square$

Theorem 3 is an interesting qualitative result. But in general, a Nash equilibrium can be arbitrarily worse than the optimal point. Hence, we give the following quantitative result. Let $r = \{r_1, \ldots, r_N\}$ be an arbitrary charging profile, and $r^{\text{ODC}} = \{r_1^{\text{ODC}}, \ldots, r_N^{\text{ODC}}\}$ be an optimal charging profile for problem ODC. Define the *sub-optimality gap $G(r)$* at r as

$$G(r) := L(r) - L\left(r^{\text{ODC}}\right) = \left\| D + \sum_n r_n \right\|^2 - \left\| D + \sum_n r_n^{\text{ODC}} \right\|^2,$$

and the *sub-optimality ratio $R(r)$* at r as

$$R(r) := \frac{G(r)}{L(r^{\text{ODC}})} = \frac{L(r) - L\left(r^{\text{ODC}}\right)}{L(r^{\text{ODC}})}.$$

Since $L\left(r^{\text{ODC}}\right)$ is the same for all optimal charging profiles r^{ODC}, $G(r)$ and $R(r)$ are defined without ambiguity. For each EV n, all its potential charging profiles are time-shifted versions of each other. Hence,

$$\|r_{n,1}\|^2 = \ldots = \|r_{n,A_n}\|^2.$$

Define $C_n := \|r_{n,1}\|^2$, then $C_n = \|r_n\|^2$ for any feasible charging profile $r_n \in \mathcal{F}_n$.

Theorem 4: *Let $r^e = \{r_1^e, \ldots, r_N^e\}$ be an arbitrary equilibrium charging profile of Algorithm DSC, then its sub-optimality gap $G(r^e)$ satisfies*

$$G(r^e) \leq 2 \sum_{n=1}^N C_n.$$

Proof: See Appendix G. $\square$

Theorem 4 gives an upper bound for the sub-optimality gap $G(r^e)$, but we are more interested in the sub-optimality ratio $R(r^e)$. Hence, we derive an upper bound $\tilde{R}$ for $R(r^e)$

$$\tilde{R} := \frac{2 \sum_n C_n}{L\left(r^{\text{ODC}}\right)} \geq R(r^e).$$

When the number N of EVs is small, the norm $\|D\|$ for non-EV demand is much larger than the norm $\left\| \sum_n r_n^{\text{ODC}} \right\|$ for aggregated EV demand. Then

$$\begin{aligned} 2 \sum_n C_n &= 2 \sum_n \left\| r_n^{\text{ODC}} \right\|^2 \leq 2 \left\| \sum_n r_n^{\text{ODC}} \right\|^2 \\ &\ll \|D\|^2 \leq L\left(r^{\text{ODC}}\right), \end{aligned}$$
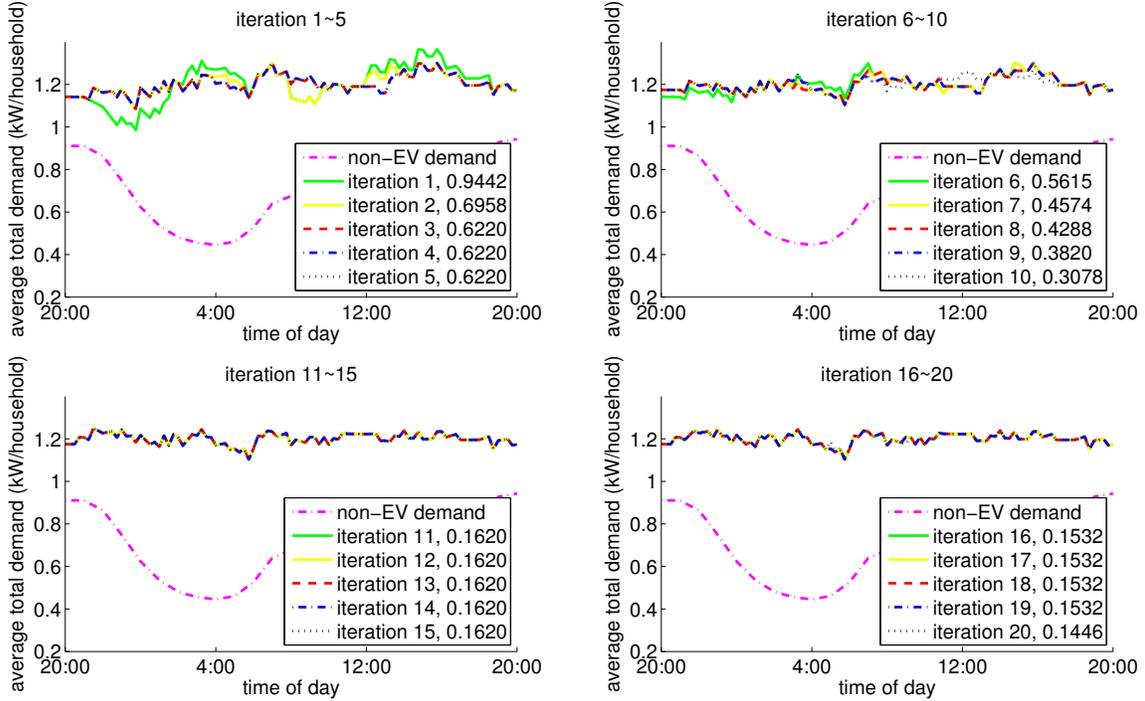
Fig. 4. Average total demand profiles (per household) and escape probabilities in the first 20 iterations of Algorithm DSC in a 100 EV case. Escape probabilities for each iteration are shown in the legends.

and the upper bound $\tilde{R}$ is much smaller than 1.

When the number N of EVs is big, the norm $\|D\|$ for non-EV demand is much smaller than the norm $\left\|\sum_n r_n^{\text{ODC}}\right\|$ for aggregated EV demand. Then we derive another upper bound

$$\hat{R} := \frac{2\sum_n C_n}{\left\|\sum_n r_n^{\text{ODC}}\right\|^2} \geq \frac{2\sum_n C_n}{\left\|D + \sum_n r_n^{\text{ODC}}\right\|^2} = \tilde{R} \geq R(r^e)$$

for the sub-optimality ratio $R(r^e)$. Note that as N increases, $\|D\|$ becomes increasingly negligible in comparison with $\left\|\sum_n r^{ODC}\right\|$, and $\hat{R}$ and $\tilde{R}$ become closer. The reasons we look into the looser upper bound $\hat{R}$ are:

(1) $\hat{R}$ does not depend on non-EV demand D, but $\tilde{R}$ does;
(2) It is clear how $\hat{R}$ scales with N. Since $2\sum_n C_n$ scales linearly with N and $\left\|\sum_n r_n^{\text{ODC}}\right\|^2$ scales quadratically with N, $\hat{R}$ scales with N as $O\left(\frac{1}{N}\right)$.

Since its upper bound $\hat{R}$ scales with N as $O\left(\frac{1}{N}\right)$, the sub-optimality ratio $G(r^e)$ remains small when N is big.

In conclusion, not only does Algorithm DSC converge to one of its equilibrium charging profiles with probability 1, but also each of its equilibrium charging profiles has a small sub-optimality ratio.

IV. CASE STUDIES

In this section, we evaluate the convergence rate and sub-optimality ratio of Algorithm DSC numerically. Consider a 24-hour scheduling horizon, and divide it into 96 slots of 15 minutes. We choose the average residential load profile in the service area of South California Edison from 20:00 on 02/13/2011 to 20:00 on 02/14/2011 [14] as the average non-EV demand per household on the scheduling horizon.

We consider different penetration levels of EVs in 100 households. Without loss of generality, we assume that all the EVs plug in at (the beginning of) slot 1 with a deadline at (the end of) slot 96. According to the EV charging characteristics in [13], we set the charging rate $r_n(t)$ to be 3.3 kW if EV n is charging at time t, and assume that each EV needs to be charged for 16 consecutive slots (4 hours).

A. Convergence Rate

Let r^k denote the charging profile of Algorithm DSC in iteration k, then the sequence $\{r^1, r^2, \ldots\}$ is a Markov chain [17, Chapter 6]. Define the escape probability $P_{\text{escape}}^k := \Pr\{r^{k+1} \neq r^k \mid r^k\}$ for $k \geq 0$.[3] If $P_{\text{escape}}^k = 0$, then r^k is an equilibrium for Algorithm DSC; otherwise, the inverse $1/P_{\text{escape}}^k$ is the expected number of iterations it takes to have a charging profile update[4]. For example, if $P_{\text{escape}}^k = 0.3$, then roughly speaking, we have to wait for 3.3 iterations before observing a charging profile update. At such an updating speed, we should stop the iterative procedure in Algorithm DSC to reduce communication overhead.

To visualize the relationship between charging profile updates and escape probability, we show the average total

[3]In fact, P_{escape}^k is the escape probability at r^k in Definition 5.
[4]If $r^{k+l-1} = \ldots = r^k$, then $\Pr\{r^{k+l} \neq r^{k+l-1} \mid r^{k+l-1}\} = p_{\text{escape}}^k$ for any $l \geq 1$. Hence, the number l of iterations it takes to have $r^{k+l} \neq r^{k+l-1}$ (charging profile update) follows a geometric distribution with parameter p_{escape}^k. Consequently, the expected number of iterations it takes to have a charging profile update is $E(l) = 1/p_{\text{escape}}^k$ [17, Chapter 3.5].

demand (per household) as well as the escape probabilities in the first 20 iterations of Algorithm DSC in a 100 EV case in Figure 4. It can be seen that the total demand profile only changes slightly from iteration 6 to iteration 10, where the escape probability is between 0.3 and 0.6; and almost does not change from iteration 11 to iteration 20, where the escape probability is below 0.3. Hence, the escape probability P_{escape}^k is a "measure" of how "close" the charging profile r^k is to an equilibrium charging profile: in this particular example, when $P_{\text{escape}}^k < 0.3$, r^k is "nearly" an equilibrium charging profile.
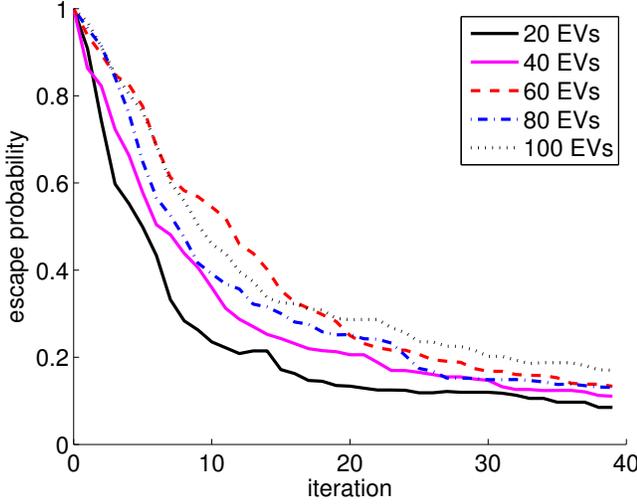

Fig. 5. Escape probability P_{escape}^k with various numbers of EVs.

Figure 5 shows the (average) escape probability P_{escape}^k (of 10 simulations) in the first 40 iterations of Algorithm DSC, with various numbers of EVs. It can be seen that the escape probability (on average) becomes less than 0.3 within 20 iterations. Hence, we terminate the iterative procedure in Algorithm DSC after 20 iterations, and let the transformer and the EVs agree on the charging profile r^{20}. We call r^{20} the output charging profile of Algorithm DSC hereafter.

B. Sub-Optimality Ratio

Figure 6 shows the average total demand (per household) in iteration 20 of Algorithm DSC, with various number of EVs. It can be seen that the total demand is always close to flat, even in the case where there are only 20 EVs. Note that completely flat total demand is not achievable since the charging rate of an EV is either 0 or 3.3kW. To quantify the sub-optimality of output charging profile r^{20}, we evaluate its sub-optimality ratio $R\left(r^{20}\right)$ numerically in this section.

Since the change in charging profiles becomes negligible after 20 iterations, we think of r^{20} as an equilibrium charging profile. Then, the sub-optimality gap $G\left(r^{20}\right)$ at r^{20} satisfies

$$G\left(r^{20}\right) \le 2\sum_{n=1}^{N} C_n = 2N\sum_{t=1}^{16} 3.3^2 = 348.48N$$

according to Theorem 4. When $G\left(r^{20}\right) \ll L\left(r^{\text{ODC}}\right)$,[5] the

[5]It can be inferred from Figure 7(a) that $G\left(r^{20}\right) \ll L\left(r^{\text{ODC}}\right)$ holds.
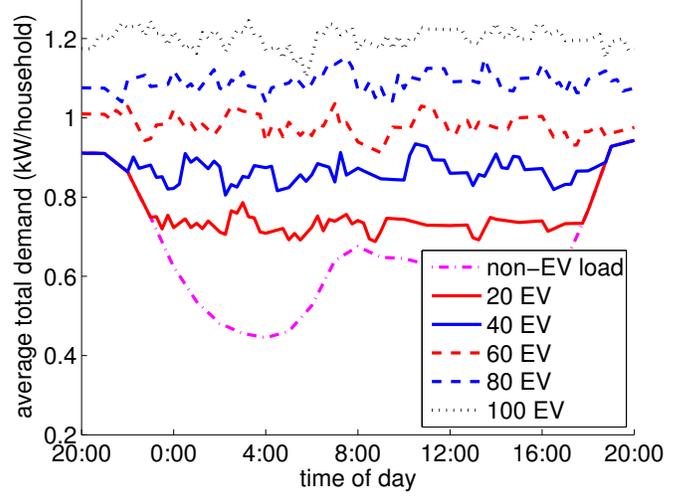

Fig. 6. Average total demand per household in iteration 20 of Algorithm DSC, with various number of EVs.

upper bound $\tilde{R}$ for sub-optimality ratio $R\left(r^{20}\right)$ satisfies

$$\tilde{R} = \frac{2\sum_n C_n}{L\left(r^{\text{ODC}}\right)} \approx \frac{348.48N}{L\left(r^{20}\right)}.$$

We plot the (approximation for) upper bound $\tilde{R}$ in Figure 7(a). It can be seen that $\tilde{R}$ is smaller than 2.6% for any number of EVs. The shape of the $\tilde{R}$ curve is explained as following:
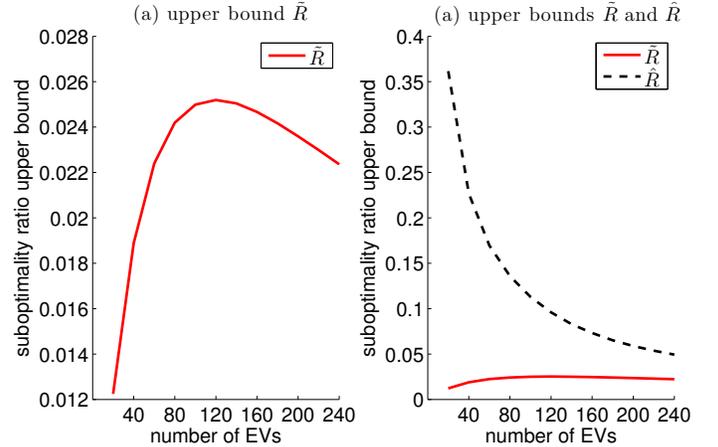

Fig. 7. Upper bounds $\tilde{R}$ and $\hat{R}$ for the sub-optimality ratio $R\left(r^{20}\right)$, as the number N of EVs varies from 20 to 240.

(a) When the number N of EVs is small, the cost $L\left(r^{20}\right) = \|D + \sum_n r_n^{20}\|^2$ is dominated by $\|D\|^2$, and $\tilde{R}$ roughly increases linearly with N.

(b) When N is big, $L\left(r^{20}\right)$ is dominated by $\|\sum_n r_n^{20}\|^2$, which scales quadratically with N. Then, $\tilde{R}$ scales with N as $O\left(\frac{1}{N}\right)$.

To emphasize (b), we plot (approximation for)[6]

$$\hat{R} = \frac{348.48N}{\|\sum_n r_n^{\text{ODC}}\|^2} \approx \frac{348.48N}{\|\sum_n r_n^{20}\|^2}$$

[6]Since $\tilde{R} \ll 1$, $L(r^{20}) \approx L\left(r^{\text{ODC}}\right)$, $\|\sum_n r_n^{\text{ODC}}\|^2 \approx \|\sum_n r_n^{20}\|^2$.

together with (approximation for) $\tilde{R}$ in figure 7(b). It can be seen that as N gets larger, $\hat{R}$ roughly scales as $O\left(\frac{1}{N}\right)$, and the two upper bounds $\tilde{R}$ and $\hat{R}$ get closer.

V. CONCLUSIONS

We have taken into account the facts that charging rate is fixed and charging should not be interrupted, and formulated EV charging scheduling as a discrete optimization problem ODC, where the objective is to flatten the total demand. We have shown that under the new charging constraints, any deterministic algorithm (that uses a common information flow pattern) cannot in general flatten the total demand.

Motivated by this fact, we proposed a stochastic distributed algorithm DSC to approximately solve problem ODC in an iterative procedure. In each iteration, the transformer receives charging profiles computed by the EVs in the previous iteration, and broadcasts the corresponding normalized total demand to the EVs; each EV generates a probability distribution over its potential charging profiles accordingly, and samples from the distribution to obtain a new charging profile.

We proved that Algorithm DSC converges almost surely to one of its equilibrium charging profiles, and each of its equilibrium charging profiles has a small sub-optimality ratio. Preliminary case studies confirm fast convergence of Algorithm DSC, and indicate less than 2.6% sub-optimality after 20 iterations, with any number of EVs.
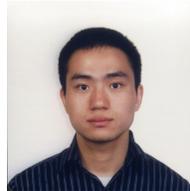
ACKNOWLEDGMENT

REFERENCES

[1] M. Duvall and E. Knipping, "Environmental assessment of plug-in hybrid electric vehicles - volume 1: nationwide greenhouse gas Emissions," Electric Power Research Institute, 2007.

[2] Committee on assessment of resource needs for full cell and hydrogen technologies, National Research Council, *Transitions to alternative transportation technologies–plug-in hybrid electric vehicles*, the National Press, 2010.

[3] U.S. Department of Energy website http://www.eere.energy.gov/. Available: http://tinyurl.com/4o5ggml

[4] L. Kelly, A. Rowe and P. Wild, "Analyzing the impacts of plug-in electric vehicles on distribution networks in British Columbia," *in Proc. Electrical Power & Energy Conference*, 2009, pp. 1-6.

[5] C. Roe, F. Evangelos, J. Meisel, A. P. Meliopoulos and T. Overbye, "Power system level impacts of PHEVs," *in Proc. Hawaii International Conference on System Science*, 2009, pp. 1-10.

[6] K. Clement, E. Haesen and J. Driesen, "Coordinated charging of multiple plug-in hybrid electric vehicles in residential distribution grids," *in Proc. Power Systems Conference and Exposition*, 2009, pp. 1-7.

[7] C. Quinn, D. Zimmerle and T. H. Bradley, "The effect of communication architecture on the availability, reliability, and economics of plug-in hybrid electric vehicle-to-grid ancillary services," *Journal of Power Sources*, vol. 195, issue 5, 2010, pp. 1500-1509.

[8] J. A. P. Lopes, P. M. R. Almeida, and A. M. M. da Silva, "Smart charging strategies for electric vehicles: Enhancing grid performance and maximizing the use of variable renewable energy sources," *in Proc. International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium and Exhibition*, 2009, pp. 1-11.

[9] Personal communication with Marshall Parsons , manager of Product Development Advanced Technology at Southern California Edison.

[10] E. Sortomme, M. M. Hindi, S. D. J. MacPherson and S. S. Venkata, "Coordinated charging of plug-in hybrid electric vehicles to minimize distribution system losses," *IEEE Transactions on Smart Grid*, vol. 2, no. 1, 2011, pp. 198-205.

[11] Z. Ma, D. Callaway and I. Hiskens, "Decentralized charging control for large populations of plug-in vehicles," *in Proc. Conference on Decision and Control*, 2010, pp. 206-212.

[12] L. Gan, U. Topcu and S. H. Low, "Optimal decentralized protocols for electric vehicle charging," *in Proc. Conference on Decision and Control*, 2011.

[13] A. Ipakchi and F. Albuyeh, "Grid of the future," *IEEE Power and Energy Magazine*, vol. 7, no. 2, 2009, pp. 52-62.

[14] South California Edison website http://www.sce.com/. Available: http://www.sce.com/005_regul_info/eca/DOMSM11.DLP

[15] S. Boyd and L. Vandenberghe, "Convex Optimization," Cambridge University Press, 2004.

[16] C. H. Papadimitriou and K. Steiglitz, "Combinatorial Optimization: algorithms and complexity," Dover Press, 1998.

[17] G. R. Grimmett and D. R. Stirzaker, "Probability and Random Process," third edition, Oxford Science Publication, 1992.

[18] P. J. M. Laarhoven and E. H. L. Aarts, "Simulated Annealing: Theory and Applications," Kluwer Academic Publishers, 1987.

[19] D. E. Goldberg, "Genetic algorithms in search, optimization, and machine learning," Addison-Wesley Publication, 1989.

Lingwen Gan received his B.S. degree from Tsinghua University, Beijing, China, in electronics engineering. He is currently working towards a Ph.D. in Electrical Engineering at California Institute of Technology, Pasadena. His main research interests are in the distributed control and optimization of networked systems such as smart grid.

Ufuk Topcu is a postdoctoral scholar in Control and Dynamical Systems at the California Institute of Technology. His research focuses on developing analytical and computational tools that aim to systematically and automatically close the specification-design-verification loop for networked, information-based system with application in autonomy, transportation vehicles, and energy networks. He received his Ph.D. in 2008 from the University of California, Berkeley.

Steven H. Low received BS from Cornell and PhD from Berkeley, both in electrical engineering. He is a Professor of the Computing & Mathematical Sciences and Electrical Engineering departments at Caltech, and hold guest faculty positions with the Swinbourne University, Australia and Shanghai Jiaotong University, China. Prior to that, he was with AT&T Bell Laboratories, Murray Hill, NJ, and the University of Melbourne, Melbourne, Australia. He was on the Editorial Board of the IEEE/ACM Transactions on Networking, IEEE Transactions on Automatic Control, ACM Computing Surveys, and the Computer Networks Journal. He is currently a Senior Editor of the IEEE Journal on Selected Areas in Communications, and series mentor for its Smart Grid Series.

APPENDIX

A. Proof of Proposition 1

For any given deterministic algorithm, let g denote its corresponding deterministic map in (3). We prove Proposition 1 by induction over the iteration index k.

- $k = 1$: For all n, the new charging profile r_n^1 is calculated according to (3) with deterministic map g. Since the inputs d^1 and S_n is the same for all n, and $r_n^0 = 0$, r_n^1 is the same for all n.

- Assume r_n^k is the same for all n in iterations $k = 1, \ldots, m$ ($m \geq 1$), now we show that r_n^{k+1} is the same for all n in iteration $m + 1$.

    For all n, the new charging profile r_n^{m+1} is calculated according to (3) with deterministic map g. Since the inputs d^{m+1}, S_n, r_n^m is the same for all n, r_n^{m+1} is the same for all n.

Hence, Proposition 1 follows. □

B. Preliminaries for Supermartingale

Constructing a supermartingale is the starting point of designing Algorithm DSC. We now provide some preliminaries following [17, chapter 12]. For a real-valued scalar x, define its negative part $x^- := \max\{0, -x\}$.

Supermartingale: Let $\mathscr{F}$ be a filtration of the probability space $(\Omega, \mathscr{F}, P)$, and let Y be a sequence of random variables which is adapted to $\mathscr{F}$. We call the pair $(Y, \mathscr{F})$ a *supermartingale* if, for all $k \geq 0$,

(a) $E(Y_k^-) < \infty$;

(b) $E(Y_{k+1} \mid \mathscr{F}_k) \leq Y_k$.

We can change $k \geq 0$ to $k \geq 1$ in the definition. The following martingale convergence theorem [17, page 454] will be used in Appendix C.

Martingale convergence theorem: Let $(Y, \mathscr{F})$ be a supermartingale and suppose that $E(Y_n^-) \leq M$ for some M and all n. These exists a random variable Y_∞ such that $Y_n \xrightarrow{a.s.} Y_\infty$ as $n \to \infty$.

C. Constructing Supermartingale

Let sample space Ω be the set of all sequences $\{r^0, r^1, \ldots\}$ of charging profiles generated by Algorithm DSC, and define σ-algebra $\mathscr{F}_k := \sigma(r^0, \ldots, r^k)$ for $k \geq 0$, then $\mathscr{F} := \{\mathscr{F}_0, \mathscr{F}_1, \ldots\}$ is a filtration. Given any initial distribution of r^0, the probability space $(\Omega, \mathscr{F}, P)$ is defined, and $\mathscr{F}$ is a filtration of the probability space. Define a random variable

$$L_k := \left\| D + \sum_n r_n^k \right\|^2,$$

and a stochastic process

$$L := \{L_1, L_2, \ldots\},$$

then L is adapted to $\mathscr{F}$, and $L_k^- = 0$ for all k.

We need an algorithm such that the sequence $L_1, L_2, \ldots$ generated by it converges with probability 1. Recalling the martingale convergence theorem, we only need to design the algorithm such that $(L, \mathscr{F})$ is a supermartingale. To make the pair $(L, \mathscr{F})$ a supermartingale, we are left to design the algorithm such that $E(L_{k+1}|\mathscr{F}_k) \leq L_k$ for $k \geq 1$. Without loss of generality, fix a $k \geq 1$, and define

$$D^k := D + \sum_n r_n^k, \quad d^k := \frac{D^k}{N}, \quad x_n := r_n^{k+1} - r_n^k$$

as the corresponding total demand, normalized total demand, and change in the charging profile of EV n. Define a scalar variance

$$Var(x|\mathscr{F}_k) := E\left(x'x|\mathscr{F}_k\right) - \|E(x|\mathscr{F}_k)\|^2,$$

then

$$
\begin{aligned}
&E(L_{k+1}|\mathscr{F}_k) - L_k \\
=\ & 2\sum_n \langle D^k, E(x_n|\mathscr{F}_k)\rangle + \sum_n Var(x_n|\mathscr{F}_k) \\
& + \sum_{m,n} \langle E(x_m|\mathscr{F}_k), E(x_n|\mathscr{F}_k)\rangle \\
\leq\ & 2\sum_n \langle D^k, E(x_n|\mathscr{F}_k)\rangle + \sum_n Var(x_n|\mathscr{F}_k) + N\sum_n \|E(x_n|\mathscr{F}_k)\|^2 \\
=\ & N\sum_n \left(2\langle d^k, E(x_n|\mathscr{F}_k)\rangle + \|E(x_n|\mathscr{F}_k)\|^2 + \frac{1}{N}Var(x_n|\mathscr{F}_k)\right).
\end{aligned}
$$

Noting that if r_n^{k+1} follows the probability distribution $p = (0, \ldots, 0, 1, 0, \ldots, 0)$ where the 1 corresponds to r_n^k, then

$$2\langle d^k, E(x_n|\mathscr{F}_k)\rangle + \|E(x_n|\mathscr{F}_k)\|^2 + \frac{1}{N}Var(x_n|\mathscr{F}_k) = 0. \tag{7}$$

Hence, optimizing over the set Θ_n of all probability distributions yields

$$\min_{p_n \in \Theta_n} 2\langle d^k, E(x_n|\mathscr{F}_k)\rangle + \|E(x_n|\mathscr{F}_k)\|^2 + \frac{1}{N}Var(x_n|\mathscr{F}_k) \leq 0. \tag{8}$$

Let p_n^* be a solution for the optimization (8). If r_n^{k+1} follows the distribution p_n^* for all n, then $E(L_{k+1}|\mathscr{F}_k) \leq L_k$.

Now we derive p_n^* for the optimization in (8). With probability distribution $p = (p_1, \ldots, p_{A_n})$,

$$E\left(r_n^{k+1}|\mathscr{F}_k\right) = R_n p,$$

and the scalar variance

$$
\begin{aligned}
Var(r_n|\mathscr{F}_k) &= \sum_a p_a r_{n,a}' r_{n,a} - \left\| \sum_a p_a r_{n,a} \right\|^2 \\
&= C_n - p' R_n' R_n p.
\end{aligned}
$$

Then

$$
\begin{aligned}
p_n^* &= \underset{p \in \Theta_n}{\operatorname{argmin}}\ 2\langle d^k, E(x_n|\mathscr{F}_k)\rangle + \|E(x_n|\mathscr{F}_k)\|^2 + \frac{1}{N}Var(x_n|\mathscr{F}_k) \\
&= \underset{p \in \Theta_n}{\operatorname{argmin}}\ \left\| d^k - r_n^k + R_n p \right\|^2 - \frac{1}{N}p' R_n' R_n p \\
&= \underset{p \in \Theta_n}{\operatorname{argmin}}\ \left\| \frac{N}{N-1}(d^k - r_n^k) + R_n p \right\|^2.
\end{aligned}
$$

This leads to Algorithm DSC, which we described in section III-C. Hence, in Algorithm DSC, $E(L_{k+1}|\mathscr{F}_k) \leq L_k$ for any fixed $k \geq 1$. The following theorem follows.

Theorem 5: *If the stochastic process L is generated by Algorithm DSC, then $(L, \mathscr{F})$ is a supermartingale.*

It follows from martingale convergence theorem that

Corollary 1: $L_k \xrightarrow{a.s.} L_\infty$ as $k \to \infty$.

Noting that Θ_n is a compact convex set, and R_n is full column rank for all n, the following lemma holds.

Lemma 1: *For any EV n and any iteration $k \geq 1$, the problem in* (8) *is strictly convex and has a unique solution, provided that $N \geq 2$.*

We only consider the case where $N \geq 2$ and $\mathcal{F}_n \neq \emptyset$ for all n in this paper.

D. Proof of Theorem 1

Define the subset $\mathcal{S}$ of charging profiles

$$\mathcal{S} := \{r \mid E\left(L_{k+1} \mid r^k = r\right) = L(r)\}.$$

Lemma 2: The set $\mathcal{S} \neq \emptyset$, and a charging profile r is an equilibrium for Algorithm DSC if and only if $r \in \mathcal{S}$.

Proof: Since the optimal charging profile r^* for problem ODC is in $\mathcal{S}$, $\mathcal{S} \neq \emptyset$.

($\Leftarrow$) At any $r = (r_1, \ldots, r_N) \in \mathcal{S}$, p_n^* computed according to (4) must satisfy (7) holds for $n \in \mathcal{N}$. Since the probability distribution $p = (0, \ldots, 0, 1, 0, \ldots, 0)$ where 1 corresponds to r_n^k satisfy (7) and p_n^* is unique, $p_n^* = p$. Hence, r is an equilibrium point for Algorithm DSC.

($\Rightarrow$) It is trivial that all equilibrium points are in $\mathcal{S}$. $\quad\square$

Theorem 1 directly follows from Lemma 2.

E. Proof of Theorem 2

Let $\mathcal{R}$ denote the set of all feasible charging profiles for problem ODC. Since $\mathcal{R}$ is finite, $\mathcal{S}$ as a subset of $\mathcal{R}$ is also finite. Define $\mathcal{X} := \{r \in \mathcal{R} \mid \text{if } r^0 = r, \; r^k \xrightarrow{a.s.} \mathcal{S}\}$ as the set of initial charging profiles that do not almost surely converge to $\mathcal{S}$, and $\mathcal{M}$ as the largest invariant set of $\mathcal{X}$.

Lemma 3: If $\mathcal{M} = \emptyset$, then $\mathcal{X} = \emptyset$.

Lemma 3 follows from the fact that $\mathcal{R}$ is finite and the transition matrix from r^k to r^{k+1} is k-independent.

Lemma 4: If $\mathcal{M} \neq \emptyset$, then $\mathcal{M} \cap \mathcal{S} \neq \emptyset$.

Because the charging profile $r_{\mathcal{M}}^* := \underset{r \in \mathcal{M}}{\arg\min} \; L(r) \in \mathcal{M} \cap \mathcal{S}$.

Lemma 5: $r^k \xrightarrow{a.s.} \mathcal{S}$ as $k \to \infty$.

Proof: Suppose not. Then $\mathcal{X} \neq \emptyset$, $\mathcal{M} \neq \emptyset$, $\mathcal{M} \cap \mathcal{S} \neq \emptyset$. However, $\mathcal{M} \subseteq \mathcal{X}$ implies that $\mathcal{M} \cap \mathcal{S} \subseteq \mathcal{X} \cap \mathcal{S} = \emptyset$, leading to a contradiction. $\quad\square$

Theorem 2 follows from Lemma 5 and the fact that $\mathcal{S}$ is discrete.

F. Proof of Theorem 3

Let $r^e = (r_1^e, \ldots, r_N^e) \in \mathcal{S}$ denote an arbitrary equilibrium for algorithm DSC, and $p^e = (p_1^e, \ldots, p_N^e)$ the corresponding distribution. Then for each n, p_n^e has only one non-zero entry, and satisfies

$$r_n^e = R_n p_n^e.$$

Equation (4) implies that

$$p_n^e = \underset{p \in \Theta_n}{\arg\min} \; \left\| \frac{N}{N-1}(d^e - r_n^e) + R_n p \right\|^2,$$

from which we obtain

$$r_n^e = \underset{r_n \in conv(\mathcal{F}_n)}{\arg\min} \; \left\| \frac{N}{N-1}(d^e - r_n^e) + r_n \right\|^2.$$

The first order optimality condition for this optimization problem implies that

$$\left\langle D + \sum_{k \neq n} r_k^e, \; r_n - r_n^e \right\rangle \geq 0 \qquad (9)$$

for any $r_n \in conv(\mathcal{F}_n)$. Define

$$J_n(r) := \left\| D + \sum_n r_n \right\|^2 - \|r_n\|^2,$$

then $J_n(r)$ is convex in r_n. Equation (9) implies that for any $\tilde{r} = \{r_1^e, \ldots, r_{n-1}^e, r_n, r_{n+1}^e, \ldots, r_N^e\}$, $r_n \in conv(\mathcal{F}_n)$, we have $J_n(\tilde{r}) \geq J_n(r^e)$. Restricting to $r_n \in \mathcal{F}_n$, $J_n(r) = L(r) - C_n$. Hence, $L(\tilde{r}) \geq L(r^e)$. $\quad\square$

G. Proof of Theorem 4

Let r^{ODC} denote the optimal charging profile for problem ODC with optimal value p^*. Let r^* an optimal charging profile for the following relaxed optimization problem (RLX)

$$\textbf{RLX} \begin{cases} \underset{r_1, \ldots, r_N}{\text{minimize}} & L(r_1, \ldots, r_N) \\ \text{subject to} & r_n \in conv(\mathcal{F}_n), \; n = 1, \ldots, N, \end{cases}$$

where the constraints $r_n \in \mathcal{F}_n$ for problem ODC is relaxed to the new constraints $r_n \in conv(\mathcal{F}_n)$. Then $L(r^*) \leq p^*$. Let r^e be an arbitrary equilibrium charging profile, D^e and D^* be the total demand profile corresponding to r^e and r^* respectively. Then $L(r^e) - L(r^*)$ is an upper-bound for the sub-optimality gap $L(r^e) - L(r^{\text{ODC}})$.

Theorem 6: The difference between total demand D^* and total demand D^e is upper-bounded by

$$\|D^* - D^e\|^2 \leq \sum_n C_n.$$

Proof: It follows from the first optimality condition for RLX that

$$\left\langle D + \sum_k r_k^*, \; r_n^e - r_n^* \right\rangle \geq 0 \qquad (10)$$

for all n. Combining inequalities (9) and (10), we get

$$\|D^* - D^e\|^2 = \left\| \sum_n r_n^* - \sum_n r_n^e \right\|^2$$

$$\leq \sum_n \langle r_n^e, r_n^e - r_n^* \rangle \leq \sum_n C_n. \qquad \square$$

We come back to the proof of Theorem 4.

$$L(r^{\text{ODC}}) - L(r^e) \geq L(r^*) - L(r^e)$$

$$= 2 \left\langle D^e, \sum_n (r_n^* - r_n^e) \right\rangle + \left\| \sum_n (r_n^* - r_n^e) \right\|^2$$

$$\geq 2 \sum_n \langle r_n^e, r_n^* - r_n^e \rangle \geq -2 \sum_n C_n. \qquad \square$$