3-2020

Managing electric vehicle charging: An exponential cone programming approach

Li CHEN

Long HE

Helen Yangfang ZHOU

manuscript

Managing Electric Vehicle Charging: An Exponential Cone Programming Approach

Li Chen
Institute of Operations Research and Analytics, National University of Singapore
chen_l@u.nus.edu

Long He
NUS Business School, National University of Singapore
longhe@nus.edu.sg

Yangfang (Helen) Zhou
Lee Kong Chian School of Business, Singapore Management University
helenzhou@smu.edu.sg

A key to the mass adoption of electric vehicles (EVs) is ease of charging, in which public charging will play an increasingly important role. We study the EV charging management of a charging service provider, which faces uncertainty in customer arrivals (e.g., arrival/departure time and charging requirements) and a tariff structure including demand charges (costs related to the highest per-period charging quantity in a finite horizon). We formulate this problem to minimize the total expected costs as a two-stage stochastic program. A common approach to solve this program, sample average approximation, suffers from its large-scale nature. Therefore, we develop an approach based on exponential cone programs, ECP-U and ECP-C for the uncapacitated and capacitated cases, respectively, which can be solved efficiently. We obtain ECP-U by leveraging the problem structure and also provide a theoretical performance guarantee. We obtain ECP-C by also using the idea from distributionally robust optimization to employ an entropic dominance ambiguity set. Based on numerical experiments with a model calibrated to EV charging data from the U.K., we demonstrate that ECP-C not only runs faster than sample average approximation (about sixty times faster for a representative capacity level) but also leads to a lower out-of-sample expected cost and the standard deviation of this cost. Our numerical results also shed light on the effect of the composition of demand charges in smoothing electricity load over time. Our methods to construct both ECP approximations could potentially be used to solve other two-stage stochastic linear programs.

Key words: stochastic programming, exponential cone programming, optimization with uncertainty, electric vehicle charging, demand charge

History: March 3, 2020.

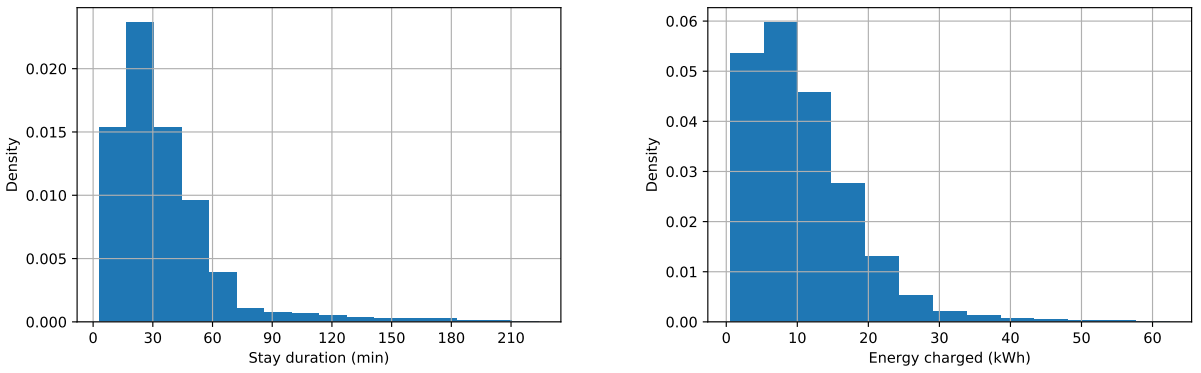
1. Introduction

Electric vehicles (EVs) are considered among the most promising technologies to decarbonize the transportation sector, with market share expected to grow from 1% in 2018 to about 30% of vehicle sales worldwide by 2030 (J.P. Morgan 2018). A key to the mass adoption of EVs is ease of charging,

where public charging will play an increasingly important role, e.g., in supporting adopters without residential charging as well as reducing "range anxiety" (Wood et al. 2017, McKinsey & Company 2018). Examples of public charging service providers include EVgo, Tesla, and ChargePoint, which have the largest market shares in the U.S. It is estimated that this charging market in the U.S. will grow to $18.6 billion and the number of charging points will expand to 40 million globally by 2030 (Bloomberg 2018).

We study the management of an EV charging service provider, which faces significant operational challenges. First, customers are stochastic in their arrival/departure time and charging requirements of their EVs: See Figure 1, which plots the density (or normalized frequency) of EV stay duration (time between arrival and departure) and energy charged using EV charging data obtained from the U.K. (U.K. Department for Transport 2018). Second, the tariff structure for an EV charging service provider in the U.S.—a commercial electricity consumer similar to operators of hospitals and office buildings—includes *demand charges*, i.e., costs "based on the highest average electricity usage occurring within a defined time interval (usually 15 minutes) during a billing period" (National Renewable Energy Laborgy 2017). The total demand charge is the sum of that applicable to all periods during the entire horizon and that applicable to periods during only on-peak or mid-peak hours. This total demand charge for an EV charging service provider can be as high as 70% of its total electricity cost (Chitkara et al. 2016).

Figure 1 **Normalized histogram of customer stay duration and energy charged using EV data from the U.K. (U.K. Department for Transport 2018)**



(a) Normalized histogram of stay duration

(b) Normalized histogram of energy charged

The above challenges call for the service provider to smooth electricity load over time in the presence of uncertainties. In this paper, we study EV charging management by a service provider

seeking to minimize the total expected cost when faced with uncertainties in customer arrival/departure time and charging requirements, and in the presence of an electricity tariff that includes demand charges. We model this problem as a two-stage stochastic program (SP), where the uncertainty is in the arrivals of customer types (defined by the arrival time, departure time, and charging requirements). In the first stage the service provider determines the charging schedule for every customer type (when the arrival of each type is uncertain). In the second stage it implements the actual charging decisions as the uncertainty is revealed over time.

Two-stage stochastic programs are in general #P-hard (Hanasusanto et al. 2016), a computational complexity class widely believed not to be polynomial-time solvable. A common approach to solve an SP is sample average approximation (SAA) based on Monte Carlo sampling (Shapiro et al. 2009, Birge and Louveaux 2011), which has been applied successfully in solving stochastic programs approximately and is asymptotically optimal when the number of samples increases. However, our SP is large-scale: For a representative EV charging problem when the planning horizon is only a day, the number of random variables (customer types) and decision variables are about 15,000 and 250,000, respectively. Thus, it is hard for SAA to provide a near-optimal solution within a reasonable amount of time, as it requires generating a sufficiently large number of samples (Shapiro and Nemirovski 2005) and the computational complexity of this approach increases usually exponentially (at least linearly) in this sample size (Kleywegt et al. 2002). Therefore, we explore another approach to solve this large-scale SP approximately.

Our proposed approach approximates our SP using exponential cone programming (Chares 2009), which can be computed efficiently (in polynomial time) owing to the recent availability of solvers, such as MOSEK (MOSEK ApS 2019). Exponential cone programming is a new class of conic programming (Boyd and Vandenberghe 2004) that generalizes linear programming to incorporate generalized inequalities defined by an exponential cone—a three-dimensional convex cone involving exponentials and logarithms. It has been used only in a limited number of papers, e.g., Chen et al. 2007, See and Sim 2010, Jaillet et al. 2018, and Zhu et al. 2019, but none existing method to derive exponential cone programs (ECPs) applies in our EV charging problem. Therefore, we develop new methods to derive our ECPs to solve our SP approximately. In addition, unlike all papers that use ECPs, which solve them using second-order cone approximations, we solve our ECPs exactly.

In order to derive an ECP when the charging service provider has a finite number of chargers, we first derive for the special case in which the provider has unlimited chargers and obtain ECP-U (where U represents "uncapacitated"). Combining the method in deriving ECP-U with an idea from distributionally robust optimization (DRO), we derive for the case with a finite number of chargers another approximation, ECP-C (where C represents "capacitated"). When theoretical performance

guarantee of our ECP approximations is hard to obtain, we conduct an extensive set of numerical experiments using a model calibrated to EV charging data from the U.K. in 2017 supplemented by data from the literature. In particular, we examine the out-of-sample performance of our ECP approach with SAA as well as a greedy policy, commonly used in practice and charging each vehicle at the maximum charging speed (Zhang et al. 2019). We also examine how the electricity load over time is affected by the composition of demand charges, a combination of all-period, on-peak, and mid-peak demand charges (applicable to periods during all hours, on-peak hours, and mid-peak hours, respectively). We summarize our contributions as follows:

- When the charging service provider has unlimited chargers, we obtain ECP-U by representing the epigraph of the moment-generating functions (MGFs) of the random arrivals of customer types using exponential cones. We show that the optimal expected cost of ECP-U gives an upper bound of that of our uncapacitated SP. We also provide a theoretical out-of-sample performance guarantee for the approximation against this SP.
- When the provider has a finite number of chargers, we obtain ECP-C by also using the idea from DRO to employ an entropic dominance ambiguity set, e.g., in Chen et al. (2019), constructed for the distributionally robust counterpart of our SP. We show that the optimal expected cost of ECP-C is an upper bound of that of our capacitated SP. We demonstrate that this expected cost of ECP-C increases as the number of chargers increases and converges to a value upper bounded by that of ECP-U when the number of chargers increases to infinity.
- We find that there is a significant value in building an SP like ours to solve the EV charging problem, as ECP-C results 11% lower expected costs than the greedy policy. Compared to SAA, our ECP-C runs much faster (about sixty times faster for a representative capacity level). In addition, ECP-C results in not only a lower expected cost but also a lower associated standard deviation, contrary to well-known tradeoff between mean cost and variability of this cost where a solution with a lower mean cost is usually less robust, i.e., a higher standard deviation of this cost (Choi and Ruszczyński 2008, Hao et al. 2019).
- We also find that either an on-peak or mid-peak demand charge alone is not sufficient to smooth electricity load, but an all-period demand charge alone can smooth electricity load as well as when all three types of demand charge exist. This is consistent with an observation in practice in a similar setting (Zhang and Qian 2018), which we show to hold when the decisions are near optimal in the EV charging management setting.

Our contributions have both practical and methodological relevances. Practically, our ECP-C approximation can be used by charging service providers in managing EV charging due to its low computation time and its superior performances. In addition, our numerical results also provide

regulatory bodies insights on designing tariff structure in smoothing electricity load. Methodologically, though originated from EV charging problems, our methods to construct ECP approximations could potentially be used to solve general two-stage stochastic linear programs if the epigraph of the MGF of the random variables can be represented using exponential cones (e.g., independent variables from common distributions such as normal, Poisson, gamma, or binomial), or if the MGF of the random variables can be upper bounded by a function whose epigraph can be characterized by exponential cones.

The rest of the paper is organized as follows. After reviewing the literature in §2, we introduce our two-stage SP in §3. In §4 we first develop ECP-U for the case in which the charging service provider has unlimited chargers, and provide its performance guarantee; we then develop ECP-C for the limited case. In §5 using a model calibrated to data, we compare the performance of ECP-C with other approaches and examine the effect of the composition of the total demand charge on electricity load. We conclude and discuss future work in §6. Any proof not in the main text can be found in the appendix.

2. Literature Review

Our paper contributes to content-related literature on EV management, as well as methodology-related literature on optimization under uncertainty.

The literature in EV operations management is nascent and focuses on problems such as the planning and operations of battery swapping stations (Mak et al. 2013, Sun et al. 2019, Schneider et al. 2017), the adoption of EVs (Avci et al. 2014, Lim et al. 2014), EV service region design (He et al. 2017), charging infrastructure planning (He et al. 2019), and fleet management of charging stations (Zhang et al. 2019). In particular, He et al. (2019) present an integrated model to jointly determine the size and location of EV charging stations, along with coupled fleet charging and repositioning operations. Zhang et al. (2019) formulate a two-stage stochastic integer program to study service-zone and facility capacity planning and fleet management in EV sharing systems with vehicle-to-grid integration. Other than the obvious differences in the problem setting, both papers assume constant charging rates in each period, which are the decision variables in our paper.

On the literature of EV charging management from the perspective of a charging service provider, Jin et al. (2013) and Zhang et al. (2014) consider the problem of scheduling EV charging in the presence of an energy storage system. Our paper differs from the first in that our EV arrivals are stochastic while theirs is deterministic and differs from the second in that we optimize the charging decisions while their charging rate is fixed. Jiang and Powell (2016) optimize the charging of one EV within a reservation window while considering the risk of the charging cost. Different from this paper, we model the setting where the provider charges multiple EVs with stochastic arrivals. In

addition, different from all these three papers, where the provider buys electricity from a market with stochastic prices, we consider the setting where the provider buys electricity from a utility company with a tariff structure that includes demand charges. This presents a unique technical challenge, and thus we develop a new method based on exponential cone programming.

We also contribute to the literature on optimization with uncertainty. Common approaches include stochastic programming (Danzig 1955, Shapiro et al. 2009, Birge and Louveaux 2011), robust optimization (e.g., Ben-Tal et al. 2009, Goldfarb and Iyengar 2003, Bertsimas and Sim 2004, Ben-Tal et al. 2013) and, more recently, distributionally robust optimization (e.g., Delage and Ye 2010, Wiesemann et al. 2014, Esfahani and Kuhn 2018, Bertsimas et al. 2019, and the references therein). We show that due to the large-scale nature of our EV charging problem, a common stochastic programming approach, SAA, is unable to provide near-optimal solutions within a reasonable amount of time. We thus develop a new approach based on ECPs, which can be solved in polynomial time owing to recent advances in the efficient computation of ECP (Dahl and Andersen 2019) and the availability of conic programming solvers (MOSEK ApS 2019). ECPs have been used in a limited number of papers, such as Chen et al. (2007), See and Sim (2010), Jaillet et al. (2018), and Zhu et al. (2019), but none existing method to derive ECPs applies to our EV charging problem. We thus develop two new ECP approximations. In particular, we develop ECP-C using the idea in the DRO framework to employ the entropic dominance ambiguity set (Chen et al. 2019). Compared to SAA, it results in not only a lower out-of-sample expected cost but also a lower standard deviation of this cost. In addition, all earlier work that use ECPs approximate them using second-order cones due to lack of efficient solvers, while we solve ECPs exactly.

Notation. We use boldface uppercase and lowercase characters to denote matrices and vectors, respectively. For example, $\boldsymbol{x} \in \mathbb{R}^n$ means that $\boldsymbol{x}$ is an n-dimensional real vector. For a matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{A}'$ is the transpose. As usual, we use $\mathbf{1}$ to denote a vector of all 1's and $\mathbf{0}$ to denote the vector or a matrix of all 0's. By $\boldsymbol{x} \geq \boldsymbol{y}$, we mean that x is greater than y component-wise. We denote by $[N] \triangleq \{1, 2, \ldots, N\}$ the set of positive running indices up to N. We use $\lfloor\cdot\rfloor$ ($\lceil\cdot\rceil$) as a floor (ceil) function that takes a real number as input and gives as output the greatest (smallest) integer less (greater) than or equal to the input. We use $\mathcal{P}_0(\mathcal{Z})$ to represent the set of all probability distributions on support set $\mathcal{Z}$. A random vector, $\tilde{\boldsymbol{z}}$, is denoted with a tilde sign and we use $\tilde{\boldsymbol{z}} \sim \mathbb{P}$, $\mathbb{P} \in \mathcal{P}_0(\mathbb{R}^{I_z})$ to define $\tilde{\boldsymbol{z}}$ as an I_z-dimensional random variable with distribution $\mathbb{P}$. We use $\mathbb{E}_{\mathbb{P}}[\cdot]$ to signify the expectation with respect to $\mathbb{P}$.

3. A stochastic programming model

We model the operations of a charging service provider to minimize the total expected cost in a finite horizon of T periods. The service provider has a limited number of chargers, denoted by C,

so customers who do not find available chargers on arrival leave the station without service. This assumption is due to the fact that the information on charger availability can be readily accessed by customers (via mobile apps), and thus customers only drive to stations with available chargers.

EV Arrival Process. EV customers differ in their arrival period, departure period, and charging requirements. We classify them into V types. We denote for each $v \in [V]$ the arrival period, departure period, and charging requirements associated with this type by $s_v \in [T], \tau_v \in [T]$, and u_v, respectively, where $s_v \le \tau_v$.[1] We assume the number of customer type v arriving at stations (with and without service) is an independent Poisson random variable with arriving rate λ_v, but due to the limited capacity of the service provider, the random number of customer type $v \in [V]$ who can find a charger on arrival (denoted by $\tilde{z}_v$) is a truncated Poisson random variable. For customers arriving in the same time t, we allow any way of truncation. Specifically, $\tilde{\boldsymbol{z}} \triangleq (\tilde{z}_v)_{v\in[V]}$ is the vector of these truncated Poisson random variables with arrival rate $\boldsymbol{\lambda} \triangleq (\lambda_v)_{v\in[V]}$ drawn from the joint distribution denoted by $\mathbb{P}^C$ (where C represents "capacitated"), as follows:

$$\sum_{v\in\mathcal{V}_t} \tilde{z}_v \le C, \quad \forall t \in [T], \tag{1}$$

where $\mathcal{V}_t \triangleq \{v \in [V] \mid s_v \le t \le \tau_v\}$ is the set of customer types at the station in period t.

Decisions. The service provider determines for each period t the amount of electricity to charge the EV of each type v, denoted by $x_{v,t}$, where $t \in \mathcal{T}_v \triangleq \{s_v, \ldots, \tau_v\}$ is the set of periods within the charging window of customer type v (i.e., the time between the arrival and departure times of this type). We fix this charging schedule $\boldsymbol{x} \triangleq (x_{v,t})_{v\in[V],t\in\mathcal{T}_v}$ at the beginning of the planning horizon.

Sequence of events. The sequence of events in each period $t \in [T]$ is as follows:

- At the beginning of period t, the charging service provider observes the number of type v's where $s_v = t$, i.e., the realization of $\tilde{z}_v$ with $s_v = t$.
- The provider charges EVs of type $v \in \mathcal{V}_t$ according to the predetermined charging schedule $\boldsymbol{x}$.
- At the end of period t, customers with departure time equal to t (i.e, $\tau_v = t$) finish their charging service and leave the charging station.

We can modify our model to represent a steady state situation, where type v with $\tau_v < s_v$ represents customers departing in period $T + \tau_v$. In particular, we can modify $\mathcal{V}_t$ to $\{v \in [V] \mid s_v \le t \le \tau_v\} \cup \{v \in [V] \mid s_v < T + t \le T + \tau_v, s_v > \tau_v\}$ (i.e., $\mathcal{V}_t$ also includes customers departing in periods after T) and modify $\mathcal{T}_v$ for customers departing in period $T + \tau_v$ to $\{s_v, \cdots, T\} \cup \{1, \cdots, \tau_v\}$. For instance, customer type v departing in period $T + 1$ has $\tau_v = 1$, and is available for charging in period 1 (i.e., $v \in \mathcal{V}_1$ and $1 \in \mathcal{T}_v$). All of our analytical results extend in a straightforward manner.

[1] An EV is not released before the departure period requested by the customer, even when the customer's charging requirements are fulfilled.

EV battery. All EV batteries have the same efficiency, denoted by $\eta \in (0,1]$, which is the ratio of the amount of electricity increased in the battery to the amount of electricity used to charge the battery. Hence, given $x_{v,t}$, the electricity in the EV battery of customer type v in period t is increased by $\eta x_{v,t}$. Since customer type v needs to fulfill the charging requirement, u_v, before τ_v, we have

$$\sum_{t \in \mathcal{T}_v} \eta x_{v,t} = u_v, \quad \forall v \in [V].$$

In addition, governed by the cell design, each EV battery has a maximum charge current or power it can accept. This constrains the maximum amount of energy a battery can be charged in each period, denoted by K (in energy unit/period), as follows:

$$0 \le x_{v,t} \le K/\eta, \quad \forall v \in [V], t \in \mathcal{T}_v.$$

We assume for expositive simplicity that the EV battery of all customers is the same. For the case in which the efficiency and maximum charging power of EV batteries differ, our theoretical results in §4 extend in a straightforward manner.

Total cost. The electricity tariff a charging service provider is subject to includes both energy charge and demand charge in a finite horizon, i.e., the total cost for the provider is the sum of these two. The energy charge is the product of the unit time-dependent electricity purchasing cost, denoted by e_t (in \$/energy unit), and the total energy purchased from the utility company in the horizon. The demand charge is the product of a unit cost, denoted by d (in \$/(energy unit/period)), and the maximum per-period total purchased electricity within the same horizon. This demand charge incentivizes the service provider to smooth the total electricity load of charging over time. If we denote by $f_t(\boldsymbol{x}, \tilde{\boldsymbol{z}})$ the total amount of electricity purchased in any period t, it is given as follows:

$$f_t(\boldsymbol{x}, \tilde{\boldsymbol{z}}) = \sum_{v \in \mathcal{V}_t} x_{v,t} \tilde{z}_v, \tag{2}$$

where the summation is over all of the customers present at the station in time t. Therefore, the demand charge is:

$$d \max_{t \in [T]} \{ f_t(\boldsymbol{x}, \tilde{\boldsymbol{z}}) \}.$$

Note that for simplicity of exposition we model the total demand charge as consisting of only all-period demand charge (applicable to all periods in the finite horizon). We consider the case in which the total demand charge is the sum of multiple types of demand charge in the numerical study in §5—for instance, including also the demand charge in peak hours and mid-peak hours.

If we denote given $(\boldsymbol{x}, \tilde{\boldsymbol{z}})$ the total cost of charging all EVs within the horizon as $c(\boldsymbol{x}, \tilde{\boldsymbol{z}})$, this is given as follows:

$$c(\boldsymbol{x}, \tilde{\boldsymbol{z}}) \triangleq \sum_{s \in [T]} e_s f_s(\boldsymbol{x}, \tilde{\boldsymbol{z}}) + d \max_{t \in [T]} \{ f_t(\boldsymbol{x}, \tilde{\boldsymbol{z}}) \}, \tag{3}$$

where the first and second terms are energy charge and demand charge, respectively.

Model formulation. The objective of the service provider is to minimize the total expected cost, where the expectation is over $\tilde{\boldsymbol{z}}$. We formulate the problem as a two-stage SP, in which we determine the *here-and-now* decisions $\boldsymbol{x}$ in the first stage and *wait-and-see* decisions for all $t \in [T]$ in the second stage, as follows:

$$\min_{x \in \mathcal{X}} \mathbb{E}_{\mathbb{P}^C}[c(\boldsymbol{x}, \tilde{\boldsymbol{z}})] \tag{4}$$

where

$$\mathcal{X} \triangleq \left\{ \boldsymbol{x} \;\middle|\; \begin{array}{ll} \sum_{t \in \mathcal{T}_v} \eta x_{v,t} = u_v & \forall v \in [V] \\ 0 \le x_{v,t} \le K/\eta & \forall v \in [V], t \in \mathcal{T}_v \end{array} \right\}.$$

Note that (4) is a stochastic linear program as the objective function can be rewritten as follows:

$$\begin{aligned} c(\boldsymbol{x}, \tilde{\boldsymbol{z}}) &= \max_{t \in [T]} \left\{ \sum_{s \in [T]} e_s f_s(\boldsymbol{x}, \tilde{\boldsymbol{z}}) + d f_t(\boldsymbol{x}, \tilde{\boldsymbol{z}}) \right\} \\ &= \min y \\ &\text{s.t. } y \ge \sum_{s \in [T]} \sum_{v \in \mathcal{V}_s} e_s x_{v,s} \tilde{z}_v + d \sum_{v \in \mathcal{V}_t} x_{v,t} \tilde{z}_v \qquad \forall t \in [T] \\ &\quad y \in \mathbb{R}. \end{aligned}$$

We denote an optimal solution to (4) by $\boldsymbol{x}^*$ and the optimal value of (4) by π^*, i.e., $\pi^* = \mathbb{E}_{\mathbb{P}^C}[c(\boldsymbol{x}^*, \tilde{\boldsymbol{z}})]$.

Challenge of solving our SP model. Two-stage SPs are in general #P-hard, as evaluating the expectation of the random objective function is #P-hard (Hanasusanto et al. 2016). In addition, our SP model (4) is large scale due to its many random variables, V, and decision variables: V is the product of three terms—the number of possible arrival periods, departure periods (given an arrival period), and charging requirements. The number of the latter the product of V and the number of periods within the charging window of all types. Consider a 1-day planning horizon: If each period is 15 minutes (consistent with the practice), then $V \approx 96 \times 16 \times 10 = 15,360$ (where 96 is the number of 15-minute arrival periods, 16 is the number of 15-minute departure periods assuming a 4-hour stay, and 10 is the number of possible charging requirement levels). Then the number of decision variables is V times 16, i.e., $\approx 15,360 \times 16 = 245,760$.

The most common approach to solve a two-stage SP is SAA (recall from §1 it is short for sample average approximation) based on Monte Carlo sampling: Generate from the true distribution many samples and replace the true distribution in the SP with the empirical distribution of these samples.

Given sample size N, we denote by $\boldsymbol{z}^i \equiv (\tilde{z}_v^i)_{v\in[V]}, \forall i \in [N]$ independently sampled from $\mathbb{P}^C$. Then the SAA model for (4) is

$$\begin{aligned} \inf_{\boldsymbol{x},\gamma^i \forall i\in[N]} \ & \frac{1}{N}\sum_{i\in[N]}\left(\sum_{s\in[T]} e_s\left(\sum_{v\in\mathcal{V}_s} x_{v,s} z_v^i\right) + d\gamma^i\right) & \\ \text{s.t.} \ & \sum_{v\in\mathcal{V}_t} x_{v,t} z_v^i \le \gamma^i & \forall t\in[T], i\in[N] \\ & \boldsymbol{x}\in\mathcal{X}. & \end{aligned} \tag{5}$$

Let us denote by $\boldsymbol{x}^S$ an optimal solution to (5) and by π^S the value of the objective function in (4) evaluated at $\boldsymbol{x}^S$, i.e., $\pi^S = \mathbb{E}_{\mathbb{P}^C}[c(\boldsymbol{x}^S, \tilde{\boldsymbol{z}})]$. Since $\boldsymbol{x}^S$ and π^S depend on the samples generated and are thus random, so to obtain a solution near optimal to (4) with a high probability (i.e., such that π^S is close to π^*) requires the sample size N to be sufficiently large (Shapiro and Nemirovski 2005), especially given the large-scale nature of (4). Furthermore, the complexity of solving (5) increases at least linearly (usually exponentially) in N (Kleywegt et al. 2002). Hence, the computation time of the SAA approach may become prohibitively long.

Therefore, it is necessary to consider other approaches that are computationally efficient. In the next section we develop tractable (polynomial time solvable) approximations based on exponential cones.

4. Exponential cone programming approximations

In this section we first discuss the uncapacitated case with an infinite number of chargers, for which we develop an ECP approximation and provide a theoretical performance guarantee. We then discuss the capacitated case and develop another ECP approximation based on the framework of DRO.

As mentioned before, an ECP is a type of a conic program useful to model constraints that involve exponentials and logarithms and can be solved in polynomial time using an interior point algorithm with an efficient solver, such as MOSEK. Note that since the most difficult step in solving (4) is computing the expectation of the largest order statistic (while the energy charge cost is linear in $\tilde{\boldsymbol{z}}$), ECPs become a natural choice for the approximation of (4) because we use an MGF (involving exponential functions) of $\tilde{\boldsymbol{z}}$ to bound this order statistic.

4.1. The uncapacitated case

Let $\mathbb{P}^\infty$ denote the joint distribution of $\tilde{\boldsymbol{z}}$ for $C=\infty$. Note then $\tilde{z}_v$'s for $v\in[V]$ are independent Poisson random variables. Given any $\boldsymbol{\theta} \triangleq (\theta_v)_{v\in[V]}$, the MGF of $\tilde{\boldsymbol{z}} \sim \mathbb{P}^\infty$ is

$$\mathbb{E}_{\mathbb{P}^\infty}\left[\exp\left(\sum_{v\in[V]}\theta_v\tilde{z}_v\right)\right] = \prod_{v\in[V]}\mathbb{E}_{\mathbb{P}^\infty}\left[\exp\left(\theta_v\tilde{z}_v\right)\right] = \prod_{v\in[V]}\exp\left(\lambda_v(e^{\theta_v}-1)\right), \tag{6}$$

where the first equality is due to the independence of $\tilde{z}_v$'s and the second equality follows from the closed-form MGF expression of a Possion random variable $\tilde{z} \sim \mathbb{P}$ with arrival rate λ, i.e., $\mathbb{E}_{\mathbb{P}}\left[e^{\theta \tilde{z}}\right] = e^{\lambda(e^{\theta}-1)}$ for any θ. Since (6) involves exponential functions, we obtain an upper bound of π^*, the optimal value of (4), using an ECP as follows:

PROPOSITION 1. *When $C = \infty$, the optimal value of the following ECP gives an upper bound of π^*:*

$$
\begin{aligned}
\inf_{\boldsymbol{x}\in\mathcal{X},\kappa,\gamma,\mu>0,\boldsymbol{\xi},\boldsymbol{\zeta}} \quad & \sum_{s\in[T]} e_s f_s(\boldsymbol{x},\boldsymbol{\lambda}) + d\,(\kappa+\gamma) & & \\
\text{s.t.} \quad & \sum_{v\in\mathcal{V}_t} x_{v,t}\lambda_v \le \gamma & \forall t\in[T] & \\
& \mu \exp\left(x_{v,t}/\mu\right) \le \xi_{v,t} & \forall t\in[T], v\in\mathcal{V}_t & \qquad \text{(ECP-U)} \\
& \mu \exp\left(\left(-\kappa + \sum_{v\in\mathcal{V}_t} \lambda_v(\xi_{v,t} - x_{v,t} - \mu)\right)/\mu\right) \le \zeta_t & \forall t\in[T] & \\
& \sum_{t\in[T]} \zeta_t \le \mu &&
\end{aligned}
$$

Proof. To obtain an upper bound of π^*, we first obtain an upper bound of $\mathbb{E}_{\mathbb{P}^\infty}\left[\max_{t\in[T]} f_t(\boldsymbol{x},\tilde{\boldsymbol{z}})\right]$:

$$
\begin{aligned}
\mathbb{E}_{\mathbb{P}^\infty}\left[\max_{t\in[T]} f_t(\boldsymbol{x},\tilde{\boldsymbol{z}})\right] =& \mathbb{E}_{\mathbb{P}^\infty}\left[\max_{t\in[T]} \left(f_t(\boldsymbol{x},\tilde{\boldsymbol{z}}) - f_t(\boldsymbol{x},\boldsymbol{\lambda}) + f_t(\boldsymbol{x},\boldsymbol{\lambda})\right)\right] \\
\le& \mathbb{E}_{\mathbb{P}^\infty}\left[\max_{t\in[T]} \left(f_t(\boldsymbol{x},\tilde{\boldsymbol{z}}) - f_t(\boldsymbol{x},\boldsymbol{\lambda})\right) + \max_{t\in[T]} f_t(\boldsymbol{x},\boldsymbol{\lambda})\right] \\
=& \mathbb{E}_{\mathbb{P}^\infty}\left[\max_{t\in[T]} (f_t(\boldsymbol{x},\tilde{\boldsymbol{z}}) - f_t(\boldsymbol{x},\boldsymbol{\lambda}))\right] + \max_{t\in[T]} f_t(\boldsymbol{x},\boldsymbol{\lambda}). \qquad (7)
\end{aligned}
$$

We then obtain an upper bound of the first term in (7) given any $\mu > 0$ as follows:

$$
\begin{aligned}
\mathbb{E}_{\mathbb{P}^\infty}\left[\max_{t\in[T]} (f_t(\boldsymbol{x},\tilde{\boldsymbol{z}}) - f_t(\boldsymbol{x},\boldsymbol{\lambda}))\right] \le& \mu \ln \mathbb{E}_{\mathbb{P}^\infty}\left[\exp\left(\max_{t\in[T]} \left(f_t(\boldsymbol{x},\tilde{\boldsymbol{z}}) - f_t(\boldsymbol{x},\boldsymbol{\lambda})\right)/\mu\right)\right] \\
\le& \mu \ln \mathbb{E}_{\mathbb{P}^\infty}\left[\sum_{t\in[T]} \exp\left(\left(f_t(\boldsymbol{x},\tilde{\boldsymbol{z}}) - f_t(\boldsymbol{x},\boldsymbol{\lambda})\right)/\mu\right)\right] \\
=& \mu \ln \sum_{t\in[T]} \mathbb{E}_{\mathbb{P}^\infty}\left[\exp\left(\left(f_t(\boldsymbol{x},\tilde{\boldsymbol{z}}) - f_t(\boldsymbol{x},\boldsymbol{\lambda})\right)/\mu\right)\right] \\
=& \mu \ln \sum_{t\in[T]} \mathbb{E}_{\mathbb{P}^\infty}\left[\exp\left(\sum_{v\in\mathcal{V}_t} \frac{x_{v,t}}{\mu}(\tilde{z}_v - \lambda_v)\right)\right] \\
=& \mu \ln \sum_{t\in[T]} \prod_{v\in\mathcal{V}_t} \mathbb{E}_{\mathbb{P}^\infty}\left[\exp\left(\frac{x_{v,t}}{\mu}\tilde{z}_v\right)\right] \exp\left(-\frac{x_{v,t}}{\mu}\lambda_v\right) \\
=& \mu \ln \sum_{t\in[T]} \exp\left(\sum_{v\in\mathcal{V}_t} \lambda_v \left(e^{x_{v,t}/\mu} - 1 - x_{v,t}/\mu\right)\right), \qquad (8)
\end{aligned}
$$

where the first line is due to the convexity of an exponential function and Jensen's inequality:

$$\exp\left(\mathbb{E}_{\mathbb{P}^\infty}\left[\max_{t\in[T]}\left(f_t(\boldsymbol{x},\tilde{\boldsymbol{z}})-f_t(\boldsymbol{x},\boldsymbol{\lambda})/\mu\right)\right]\right)\le\mathbb{E}_{\mathbb{P}^\infty}\left[\exp\left(\max_{t\in[T]}\left(f_t(\boldsymbol{x},\tilde{\boldsymbol{z}})-f_t(\boldsymbol{x},\boldsymbol{\lambda})/\mu\right)\right)\right];$$

the fourth line results from substituting (2) into the right-hand side (RHS) of the third line; and the fifth and sixth lines follow from (6).

Note if we combine (7) and (8) as follows

$$\mathbb{E}_{\mathbb{P}^\infty}\left[\max_{t\in[T]}f_t(\boldsymbol{x},\tilde{\boldsymbol{z}})\right]\le\inf_{\mu>0}\left(\mu\ln\sum_{t\in[T]}\exp\left(\sum_{v\in\mathcal{V}_t}\lambda_v\left(e^{x_{v,t}/\mu}-1-x_{v,t}/\mu\right)\right)\right)+\max_{t\in[T]}f_t(\boldsymbol{x},\boldsymbol{\lambda}),$$

its RHS has the following epigraph form

$$\begin{aligned}\inf_{\mu>0,\kappa,\gamma}\ & \kappa+\gamma\\ \text{s.t.}\ & \sum_{v\in\mathcal{V}_t}x_{v,t}\lambda_v\le\gamma,\quad\forall t\in[T]\\ & \mu\ln\sum_{t\in[T]}\exp\left(\sum_{v\in\mathcal{V}_t}\lambda_v\left(e^{x_{v,t}/\mu}-1-x_{v,t}/\mu\right)\right)\le\kappa.\end{aligned}\tag{9}$$

Note also that the second constraint in (9) is equivalent to

$$\sum_{t\in[T]}\mu\exp\left(\sum_{v\in\mathcal{V}_t}\lambda_v\left(e^{x_{v,t}/\mu}-x_{v,t}/\mu-1\right)-\kappa/\mu\right)\le\mu,$$

where the expression within the first summation of the left-hand side, i.e., $\mu\exp\left(\sum_{v\in\mathcal{V}_t}\lambda_v\left(e^{x_{v,t}/\mu}-x_{v,t}/\mu-1\right)-\kappa/\mu\right)$, has an epigraph form for all $t\in[T]$ as follows:

$$\begin{aligned}\inf_{\zeta_t,\xi_{v,t}}\ & \zeta_t\\ & \mu\exp\left(\sum_{v\in\mathcal{V}_t}\lambda_v\left(\xi_{v,t}/\mu-x_{v,t}/\mu-1\right)-\kappa/\mu\right)\le\zeta_t\\ & \mu e^{x_{v,t}/\mu}\le\xi_{v,t}\qquad\qquad\forall v\in\mathcal{V}_t\end{aligned}\tag{10}$$

Hence, (9) can be equivalently written as

$$\begin{aligned}\inf_{\mu>0,\kappa,\gamma,\boldsymbol{\xi},\boldsymbol{\zeta}}\ & \kappa+\gamma\\ \text{s.t.}\ & \sum_{v\in\mathcal{V}_t}x_{v,t}\lambda_v\le\gamma & \forall t\in[T]\\ & \mu e^{x_{v,t}/\mu}\le\xi_{v,t} & \forall t\in[T],v\in\mathcal{V}_t\\ & \mu\exp\left(\left(-\kappa+\sum_{v\in\mathcal{V}_t}\lambda_v(\xi_{v,t}-x_{v,t}-\mu)\right)/\mu\right)\le\zeta_t & \forall t\in[T]\\ & \sum_{t\in[T]}\zeta_t\le\mu.\end{aligned}\tag{11}$$

Since $f_s(\boldsymbol{x}, \tilde{\boldsymbol{z}})$ is linear in $\tilde{\boldsymbol{z}}$ for all $s \in [T]$ and $\tilde{\boldsymbol{z}} \sim \mathbb{P}^\infty$ is the vector of independent Poisson random variables, we have

$$\mathbb{E}_{\mathbb{P}^\infty}\left[f_s(\boldsymbol{x}, \tilde{\boldsymbol{z}})\right] = f_s(\boldsymbol{x}, \boldsymbol{\lambda}). \tag{12}$$

Therefore, by combining (7), (8), (11), and (12) and optimizing over $\boldsymbol{x}$, we get the optimal value of ECP-U as an upper bound of π^*, as a solution optimal to ECP-U is also feasible to (4). □

The model in Proposition 1 is labeled ECP-U (recall U represents "uncapacitated") because all of the constraints in this model involve either linear functions or exponential functions: The latter constraints can be written as exponential cone constraints that involve exponential cone $\mathcal{K}_{\exp}$, which is defined as

$$\mathcal{K}_{\exp} \triangleq \{(x_1, x_2, x_3) \mid x_1 \geq x_2 \exp(x_3/x_2), x_2 > 0\} \cup \{(x_1, 0, x_3) \mid x_1 \geq 0, x_3 \leq 0\}.$$

For example, the constraint $\mu \exp\left(x_{v,t}/\mu\right) \leq \xi_{v,t}$ can be written as $(\xi_{v,t}, \mu, x_{v,t}) \in \mathcal{K}_{\exp}$.

The number of decision variables of ECP-U is on the same order as that of our SP model, i.e., (4)), though ECP-U introduces extra decision variables κ, γ, μ, $\boldsymbol{\xi}$, and $\boldsymbol{\zeta}$, where the first three are scalars, and the last two have at most the same dimensions as $\boldsymbol{x}$. Similarly, the number of constraints in ECP-U is on the same order of that of (4), i.e., for every $t \in [T]$ and $v \in \mathcal{V}_t$, though the constraints of (4) are linear and those of ECP are nonlinear. Different from the SAA formulation of (4), which requires generating a large number of samples and the computation time may become prohibitively long, ECP-U can be solved via a state-of-the-art conic programming solver efficiently (MOSEK ApS 2019).

We denote an optimal solution to ECP-U by $\bar{\boldsymbol{x}}$ and the value of the objective function of (4) evaluated at $\bar{\boldsymbol{x}}$ by π^E, i.e., $\pi^E = \mathbb{E}_{\mathbb{P}^\infty}\left[c(\bar{\boldsymbol{x}}, \tilde{\boldsymbol{z}})\right]$. We next show that compared with an optimal solution to (4), $\boldsymbol{x}^*$, there is a theoretical guarantee on the out-of-sample performance of $\bar{\boldsymbol{x}}$. Theorem 1 compares π^E and π^* (note that $\pi^* = \mathbb{E}_{\mathbb{P}^\infty}\left[c(\boldsymbol{x}^*, \tilde{\boldsymbol{z}})\right] = \mathbb{E}_{\mathbb{P}^C}[c(\boldsymbol{x}^*, \tilde{\boldsymbol{z}})]$, as $\mathbb{P}^C \equiv \mathbb{P}^\infty$ when C is infinite). Let $\phi(x) \triangleq e^x - 1 - x$.

THEOREM 1. *Any optimal solution $\bar{\boldsymbol{x}}$ to ECP-U has an out-of-sample performance guarantee as follows:*

$$\pi^E - \pi^* = \mathbb{E}_{\mathbb{P}^\infty}\left[c(\bar{\boldsymbol{x}}, \tilde{\boldsymbol{z}})\right] - \mathbb{E}_{\mathbb{P}^\infty}\left[c(\boldsymbol{x}^*, \tilde{\boldsymbol{z}})\right] \leq d \inf_{\mu > 0}\left(\sum_{v \in [V]} \lambda_v \mu \phi\left(\frac{u_v}{\eta\mu}\right) + \mu \ln(2T)\right). \tag{13}$$

Moreover, by taking $\mu := \sqrt{\sum_{v\in[V]}(\lambda_v u_v^2)/(2\ln(2T))}/\eta$, the RHS of (13) is no greater than

$$\left(\sqrt{2} + \frac{1}{\sqrt{2}}\right) \frac{d}{\eta} \sqrt{\ln(2T) \sum_{v \in [V]} \lambda_v u_v^2} \tag{14}$$

if for all $v \in [V]$

$$2u_v^2 \ln(2T) \le \sum_{v\in[V]} \lambda_v u_v^2. \tag{15}$$

Theorem 1 gives two upper bounds in (13) and (14) on the difference between π^E and π^*. The first bound takes the inf among all $\mu > 0$, and is simplified to the second bound by taking a specific value for μ in the first bound. Both bounds depend on d, η, and T, as well as λ_v and u_v (for all $v \in [V]$), but they do not depend on e_s and K.

Compared with the second bound, the first bound is more general as it does not require condition (15), while the second one does. However, condition (15) holds in all of the numerical instances (calibrated to data) considered in §5. On the other hand, compared with the first bound, the second one is more intuitive: It is easy to show that $\sum_{v\in[V]} \lambda_v u_v^2$ in (14) is the variance of $\sum_{v\in[V]} \tilde{z}_v u_v$, which is the total charging requirements over all types. Hence, the second bound is the standard deviation of this total charging requirements multiplied by a constant.

4.2. The capacitated case

Unlike the uncapacitated case, $\tilde{\boldsymbol{z}} \sim \mathbb{P}^C$ does not follow the joint distribution of independent Poisson random variables when $C < \infty$, so it is difficult to obtain its analytical characterization, similar to that in (6). We then use the idea of DRO. Instead of assuming an exact distribution, the DRO framework considers a set of distributions from an ambiguity set (denoted by $\mathcal{F}$) which contains $\tilde{\boldsymbol{z}} \sim \mathbb{P}^C$, i.e., $\mathbb{P}^C \in \mathcal{F}$. It then optimizes for robust decisions aiming to hedge against the worst-case distribution among all possible distributions in the ambiguity set. In our application, we consider the following infinitely constrained "entropic dominance" ambiguity set, adapted from Chen et al. (2019):

$$\mathcal{F} \triangleq \left\{ \mathbb{P} \in \mathcal{P}_0\left(\mathbb{R}^V\right) \;\middle|\; \begin{array}{l} \tilde{\boldsymbol{z}} \sim \mathbb{P} \\ \ln \mathbb{E}_{\mathbb{P}}\left[\exp\left(\boldsymbol{\theta}'\tilde{\boldsymbol{z}}\right)\right] \le \sum_{v\in[V]} \lambda_v \left(e^{\theta_v} - 1\right) \quad \forall \boldsymbol{\theta} \ge \mathbf{0} \\ \mathbb{P}\left[\tilde{\boldsymbol{z}} \in \mathcal{Z}\right] = 1 \end{array} \right\}, \tag{16}$$

where

$$\mathcal{Z} \triangleq \left\{ \boldsymbol{z} \ge \mathbf{0} \mid \sum_{v\in\mathcal{V}_t} z_v \le C,\ \forall t \in [T] \right\}. \tag{17}$$

The second line in (16) ensures that the random variables $\tilde{z}_v$'s are truncated from the independent Poisson variables (i.e., in uncapacitated case): The left- and right-hand sides of the second line are the log of the MGF of $\tilde{\boldsymbol{z}} \sim \mathbb{P}^C$ and $\tilde{\boldsymbol{z}} \sim \mathbb{P}^\infty$ (see (6)), respectively. Next, we show that the ambiguity set (16) contains the underlying distribution $\mathbb{P}^C$.

Proposition 2. $\mathbb{P}^C \in \mathcal{F}$.

Proof. The support constraints of $\mathbb{P}^C$ in (1) are satisfied in (17), i.e., in the definition of $\mathcal{F}$. Let $\tilde{\boldsymbol{z}} \sim \mathbb{P}^C$ and $\tilde{\boldsymbol{w}} \sim \mathbb{P}^\infty$ (i.e., the vector of independent Poisson random variables $\tilde{w}_v$ with arrival rate $\lambda_v > 0$), and denote the joint distribution of $(\tilde{\boldsymbol{z}}, \tilde{\boldsymbol{w}})$ as $\bar{\mathbb{P}}$. By the definition of $\tilde{\boldsymbol{z}}$, we know $\tilde{\boldsymbol{z}} \le \tilde{\boldsymbol{w}}$ holds almost surely. Then $\boldsymbol{\theta}'\tilde{\boldsymbol{z}} \le \boldsymbol{\theta}'\tilde{\boldsymbol{w}}$ holds for all $\boldsymbol{\theta} \ge \mathbf{0}$ with probability one. So we have

$$\ln \mathbb{E}_{\mathbb{P}^C}\left[\exp\left(\boldsymbol{\theta}'\tilde{\boldsymbol{z}}\right)\right] = \ln \mathbb{E}_{\bar{\mathbb{P}}}\left[\exp\left(\boldsymbol{\theta}'\tilde{\boldsymbol{z}}\right)\right] \le \ln \mathbb{E}_{\bar{\mathbb{P}}}[\exp\left(\boldsymbol{\theta}'\tilde{\boldsymbol{w}}\right)] = \ln \mathbb{E}_{\mathbb{P}^\infty}\left[\exp\left(\boldsymbol{\theta}'\tilde{\boldsymbol{w}}\right)\right] = \sum_{v\in[V]} \lambda_v(e^{\theta_v} - 1),$$

where the inequality is due to the monotonicity of $\ln \mathbb{E}_{\bar{\mathbb{P}}}[\exp(\cdot)]$ and the third equality follows from (6). Hence $\mathbb{P}^C \in \mathcal{F}$. □

Therefore, we can obtain an upper bound of the optimal value of (4) by considering the worst-case optimization over the ambiguity set $\mathcal{F}$:

$$\min_{\boldsymbol{x}\in\mathcal{X}} \sup_{\mathbb{P}\in\mathcal{F}} \mathbb{E}_{\mathbb{P}}\left[c(\boldsymbol{x}, \tilde{\boldsymbol{z}})\right], \tag{DRO-Ent}$$

where "Ent" refers to the "entropic dominance" ambiguity set used. However, due to the infinite number of constraints in $\mathcal{F}$ (since the second constraint defining $\mathcal{F}$ has to hold for all $\boldsymbol{\theta} \ge \mathbf{0}$), the DRO-Ent model is not tractable (Chen et al. 2019). Therefore, we approximate DRO-Ent using an ECP formulation similar to that in Proposition 1, which can be also solved efficiently. To this end, we write $\mathcal{F}$ as the intersection of the following two sets (as shown in Lemma 1):

$$\mathcal{F}^1 \triangleq \left\{ \mathbb{P} \in \mathcal{P}_0\left(\mathbb{R}^V\right) \;\middle|\; \begin{array}{l} \tilde{\boldsymbol{z}} \sim \mathbb{P} \\ \ln \mathbb{E}_{\mathbb{P}}\left[\exp\left(\boldsymbol{\theta}'\tilde{\boldsymbol{z}}\right)\right] \le \sum\limits_{v\in[V]} \lambda_v\left(e^{\theta_v} - 1\right) \quad \forall \boldsymbol{\theta} \ge \mathbf{0} \end{array} \right\} \tag{18}$$

and

$$\mathcal{F}^2 \triangleq \left\{ \mathbb{P} \in \mathcal{P}_0\left(\mathbb{R}^V\right) \;\middle|\; \begin{array}{l} \tilde{\boldsymbol{z}} \sim \mathbb{P} \\ \mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{z}}\right] \le \boldsymbol{\lambda} \\ \mathbb{P}\left[\tilde{\boldsymbol{z}} \in \mathcal{Z}\right] = 1 \end{array} \right\}, \tag{19}$$

where recall that $\boldsymbol{\lambda} \triangleq (\lambda_v)_{v\in[V]}$.

Lemma 1. $\mathcal{F} = \mathcal{F}^1 \cap \mathcal{F}^2$.

$\mathcal{F}^2$ is an example of the moment-based ambiguity sets commonly used in DRO, which usually involve at most the first two moments. Lemma 1 implies that $\mathcal{F}$ is a subset of $\mathcal{F}^2$, i.e., any $\tilde{\boldsymbol{z}}$ satisfying the second and third constraints in $\mathcal{F}$ also satisfies the constraint on the first moment in $\mathcal{F}^2$, i.e., $\mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{z}}\right] \le \boldsymbol{\lambda}$. In other words, $\mathcal{F}$ incorporates more moment information than $\mathcal{F}^2$ does, and, therefore, $\mathcal{F}$ is smaller and less conservative than $\mathcal{F}^2$.

Based on this lemma, we then write DRO-Ent as the optimization over ambiguity set $\mathcal{F}^1$ and $\mathcal{F}^2$ separately. For the former case, we can use a technique similar to that of the proof of Proposition 1;

for the latter, as $\mathcal{F}^2$ is an example of moment-based ambiguity sets used in DRO, we use techniques standard in DRO to obtain bounds. We combine these two cases to obtain a tractable upper bound, as follows:

PROPOSITION 3. *When $C<\infty$, the following ECP gives an upper bound of π^*:*

$$
\begin{aligned}
\inf_{\substack{\boldsymbol{x}\in\mathcal{X},a,\boldsymbol{b},\boldsymbol{\nu},\boldsymbol{y},\mu,\\ \kappa,\gamma,\alpha,\boldsymbol{\beta},\boldsymbol{\xi},\boldsymbol{\zeta},\boldsymbol{\rho}}} \quad & a+\boldsymbol{b}'\boldsymbol{\lambda}+d\,(\kappa+\gamma+\alpha+\boldsymbol{\beta}'\boldsymbol{\lambda}) \\
\text{s.t.} \quad & \sum_{v\in\mathcal{V}_t} y_{v,t}\lambda_v \le \gamma && \forall t\in[T] \\
& (\xi_{v,t},\mu,y_{v,t})\in\mathcal{K}_{\exp} && \forall t\in[T], v\in\mathcal{V}_t \\
& \left(\zeta_t,\mu,-\kappa+\sum_{v\in\mathcal{V}_t}\lambda_v(\xi_{v,t}-y_{v,t}-\mu)\right)\in\mathcal{K}_{\exp} && \forall t\in[T] \\
& \sum_{t\in[T]}\zeta_t\le\mu && \text{(ECP-C)} \\
& C\sum_{k\in[T]}\rho_t^k\le\alpha && \forall t\in[T] \\
& x_{v,t}-y_{v,t}-\beta_v\le\sum_{k\in\mathcal{T}_v}\rho_t^k && \forall t\in[T], v\in\mathcal{V}_t \\
& C\sum_{t\in[T]}\nu_t\le a \\
& \sum_{s\in\mathcal{T}_v}x_{v,s}e_s-b_v\le\sum_{t\in\mathcal{T}_v}\nu_t && \forall v\in[V] \\
& \boldsymbol{b}\ge\boldsymbol{0},\boldsymbol{\nu}\ge\boldsymbol{0},\boldsymbol{y}\ge\boldsymbol{0},\boldsymbol{\beta}\ge\boldsymbol{0},\boldsymbol{\rho}\ge\boldsymbol{0},\mu>0.
\end{aligned}
$$

The model in Proposition 3 is labeled ECP-C (recall C represents "capacitated"). Similar to ECP-U, compared to our SP model in (4), ECP-C introduces extra decision variables and constraints. The numbers of both are on the same order of magnitude as our SP model. Therefore, similar to ECP-U, ECP-C can be solved efficiently.

We next show the monotonicity of the optimal value of ECP-C as C increases and the connection between ECP-C and ECP-U.

THEOREM 2. *The optimal value of ECP-C increases as C increases and converges to a value that is upper bounded by the optimal value of ECP-U as $C\to\infty$.*

Proof. We denote ECP-C for the case when capacity equals C_1 and C_2 ($C_1\le C_2$) by ECP-C1 and ECP-C2, respectively. Any feasible solution to ECP-C2 is also feasible to ECP-C1: See constraints involving C in ECP-C. Hence, compared to ECP-C1, the feasible region for ECP-C2 is no larger. Since the objective function of ECP-C does not depend on C, the optimal value of ECP-C2 is higher than that of ECP-C1.

Next we prove the optimal value of ECP-C is less than that of ECP-U for any given value of C. We substitute $a=0$, $b_v=\sum_{s\in\mathcal{T}_v} x_{v,s}e_s$, $\boldsymbol{\nu}=\mathbf{0}$, $\boldsymbol{y}=\boldsymbol{x}$, $\boldsymbol{\rho}=\mathbf{0}$, $\alpha=0$, and $\boldsymbol{\beta}=\mathbf{0}$ into ECP-C: i) all the constraints are degenerated to the constraints in ECP-U; ii) the objective function becomes

$$a+\boldsymbol{b}'\boldsymbol{\lambda}+d\,(\kappa+\gamma+\alpha+\boldsymbol{\beta}'\boldsymbol{\lambda})=\sum_{s\in[T]} e_s\sum_{v\in\mathcal{V}_s} x_{v,t}\lambda_v+d\,(\kappa+\gamma)=\sum_{s\in[T]} e_s f_s(\boldsymbol{x},\boldsymbol{\lambda})+d\,(\kappa+\gamma)\,,$$

which is the same as that of ECP-U. This means that there always exists a feasible solution in ECP-C that attains the optimal value of ECP-U. Therefore, the optimal value of ECP-C for any C is upper bounded by that of ECP-U.

Since the optimal value of ECP-C is monotonic in C and bounded by the optimal value of ECP-U, it converges to a value upper bounded by the optimal value of ECP-U as $C\to\infty$. □

Theorem 2 shows that the optimal value of ECP-C increases as charging station capacity increases, intuitively because more vehicles can be admitted to the station to be charged. It also shows that this optimal value of ECP-C converges and yields an asymptotic lower bound of the optimal value of ECP-U.

However, unlike ECP-U, we are unable to obtain a theoretical performance guarantee of an optimal action to ECP-C, and thus examine that in the numerical study.

Remark 1. As seen, our method to construct ECP-U could potentially be used to solve two-stage stochastic linear programs if the epigraph of the MGF of the random variables in the SP can be represented using exponential cones. Other than independent Poisson random variables, examples of such random variables include independent variables from common distributions such as normal, gamma, and binomial. In addition, our method to construct ECP-C could potentially be used to solve two-stage stochastic linear programs if the MGF of the random variables can be upper bounded by a function whose epigraph can be characterized by exponential cones.

5. Numerical study

In this section we compare the performance of ECP-C with SAA and a greedy policy using a model calibrated to data. We also use ECP-C to study the effect of the composition of demand charge on the electricity load over time.

5.1. Numerical setup

We consider the operations of a charging service provider in a representative 1-day horizon within a typical 1-month billing cycle. Consistent with the practice, each period represents 15 minutes, length of a period for most utility companies (such as Southern California Edison), which provides electricity to commercial electricity users (such as charging service providers). Therefore, the length of the planning horizon, T, equals 96.

Electricity tariff calibration. We use Southern California Edison's electricity tariff *Schedule GS-2 General Service* (Neubauer and Simpson 2015), which is the tariff an electricity user such as an EV charging service provider is subject to. Under this tariff, the unit energy charge depends on the hour of the day as follows:

$$\hat{e}_t = \begin{cases} \$0.1466/\text{kWh} & \text{if } 13 \le \lceil t/4 \rceil \le 18 \text{ (on-peak hours)} \\ \$0.0895/\text{kWh} & \text{if } 9 \le \lceil t/4 \rceil \le 12 \text{ or } 19 \le \lceil t/4 \rceil \le 23 \text{ (mid-peak hours)} \\ \$0.0582/\text{kWh} & \text{otherwise (off-peak hours).} \end{cases}$$

where $\lceil t/4 \rceil$ represents the hour index of period t (as 1 hour has four periods). The total demand charge is the sum of three components: all-period demand charge (termed "facility-related demand charge" in the tariff structure, and applicable to all periods throughout the horizon at \$13.94/kW); peak-hour demand charge (applicable only to on-peak hours at \$16.20/kW); and mid-peak demand charge (applicable only to mid-peak hours at \$4.95/kW). We divide the three unit demand charges in a month by 30 to scale them to those of a day and obtain the corresponding $\hat{d}_t$ as follows:

$$\hat{d}_t = \begin{cases} \$0.465/\text{kW} & \forall t \in [T] \doteq [96] \\ \$0.540/\text{kW} & \text{if } 13 \le \lceil t/4 \rceil \le 18 \text{ (on-peak hours)} \\ \$0.165/\text{kW} & \text{if } 9 \le \lceil t/4 \rceil \le 12 \text{ or } 19 \le \lceil t/4 \rceil \le 23 \text{ (mid-peak hours).} \end{cases}$$

Note that though the model in §4 has only all-period demand charge for the expositive simplicity, we consider two more types of demand charges in the total cost in the numerical study. The corresponding ECP model is constructed analogously to those in §4.

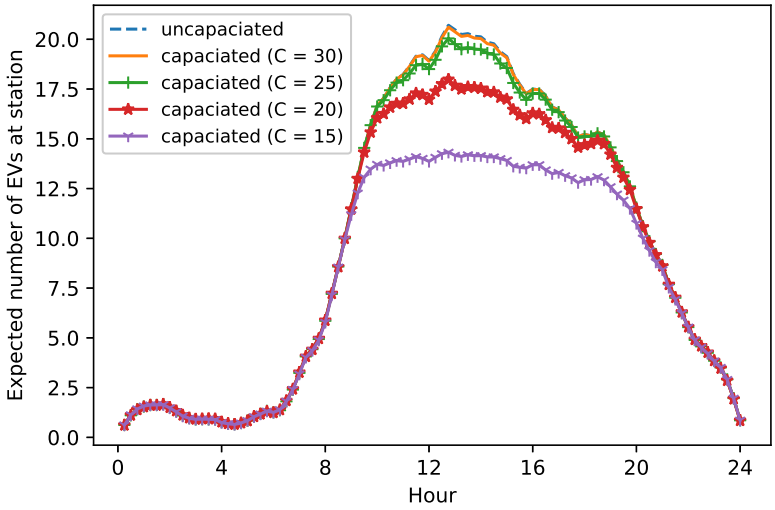
EV arrival data and their calibration. We obtain the EV charging data used for plotting Figure 1 from the U.K. Department for Transportation, which contains all of the charging events at rapid chargers (where the charging current is above 22 kV) from 27 local authorities in England from January 1 to December 31, 2017 (U.K. Department for Transport 2018). The dataset contains for each charging event the start time, end time, and energy supplied.[2] We first preprocess the data as follows: We remove those charging events where either the plug-in duration (the time interval between start time and end time) is more than 4 hours (as most are less than 4 hours), or the charging power (total energy charged divided by plug-in duration) is larger than 43 kW, or the total energy charged is either less than 0.5 kWh or greater than 62.5 kWh. In the end, we retain 93.53% of the original charging events and obtain 101,707 charging events in total, i.e., 278.65 daily charging events on average.

Discretizing time into 15-minutes intervals, we obtain the data for Figure 1: both Figure 1(a) and Figure 1(b) show the right-skewness and the long tails of the distributions of the measure of interest. The mean (standard deviation) of stay duration is 35.83 (27.92) minutes; the mean (standard deviation) of energy charged is 10.69 (7.64) kWh.

[2] To our knowledge, no such detailed EV charging data from the U.S. is publicly available.

We use these preprocessed data to estimate the parameters of the EV arrival process. The set of EV types $[V]$ is the set of distinct start times, departure times, and total energy charged, and we obtain 14,284 types. We then estimate the arrival rate of each type as the total number of charging events *of each type* within all days divided by the number of days. We plot in Figure 2 the expected number of EVs at the charging stations in a 1-day horizon when the number of chargers, C, takes different values. As can be seen, when C is infinite, the expected number of customers at the station follows a daily pattern: It starts low from midnight to 6 am, increases significantly to the peak at around 1 pm, and finally decreases significantly until midnight. When $C = 30$, there is almost no lost demand (with service level 99.86%): The corresponding line in Figure 2 almost coincides with that for the uncapacitated case. But when $C = 25, 20$, and 15, the service level drops quickly, to 98.83%, 94.07%, and 82.54%, respectively.

Figure 2 **Expected number of EVs at the station with different number of chargers (C)**



We next estimate the parameters of EVs using the parameters of Nissan Leaf, one of the most common EV types in the U.K. Specifically, we set $\hat{U} = 62$ kWh as the energy capacity of Nissan Leaf (Nissan 2019); $\hat{K} = 10.75$ kWh/period based on the rapid AC chargers with power at 43 kW (Zap-Map 2019); and $\hat{\eta} = 0.9$ as the efficiency of lithium-ion battery in Nissan Leaf (Karlsson and Kushnir 2013).

Computation. We conduct our numerical experiments in Python with the MOSEK 9.0 solver (MOSEK ApS 2019) on a Windows-OS 64-bit laptop with 24 GB of RAM and an Intel i7-8750H CPU@2.20GHz Processor. All of the problem instances are solved via a homogeneous primal-dual

interior point method with both a primal and dual feasibility tolerance of 10^{-8} and a relative duality gap of 10^{-8}.

5.2. Performance of different approaches

We compare the performance of ECP-C and SAA. We benchmark both using a greedy policy, which charges each vehicle using maximal charging speed (charging power capacity) until completion on a first-come-first-served basis and is currently used in practice and literature (Zhang et al. 2019). Denote by $\boldsymbol{x}^G$ the decision variable under the greedy policy. Then for all $v \in [V]$ and $t \in \mathcal{T}_v$, we have $x^G_{v,t} = \min\{K, u_v - (t - s_v) \cdot K\}$.

Specifically, we first compute, under each of the three approaches, the following out-of-sample quantities: (i) the mean and standard deviation of the total cost, i.e., the objective function in our SP model (4) evaluated at the solution under each approach, (ii) the average of the maximum electricity load during the entire horizon, on-peak hours, mid-peak hours, and off-peak hours, and (iii) the average of the ratio of the demand charge to the mean total cost. Using (i) as an example, after solving for $\boldsymbol{x}$ with the calibrated parameters under any approach (i.e., obtaining $\boldsymbol{x}^E, \boldsymbol{x}^S$, and $\boldsymbol{x}^G$ under SAA, ECP-C, and the greedy policy, respectively), we generate 10,000 samples via Monte Carlo simulation, implement $\boldsymbol{x}$ on these samples, and compute the sample mean and standard deviation of $c(\boldsymbol{x}, \tilde{\boldsymbol{z}})$. We refer to these sample means and sample standard deviations as the means and standard deviations unless there is ambiguity. We similarly compute (ii) and (iii). We also compare the computation time for ECP-C and SAA (as the greedy policy can be computed instantaneously): The computation time for SAA is taken as the average computation time of five numerical instances, as SAA is a random approximation (while ECP-C is a deterministic approximation).

Table 1 lists the performance measures (i), (ii), and (iii) of ECP-C and SAA relative to that of the greedy policy in columns 3-4, columns 5-8, and column 9, respectively. Using column 3 as an example, if we denote the objective function in (4) evaluated at the greedy policy by π^G (i.e., $\pi^G = \mathbb{E}_{\mathbb{P}^C}[c(\boldsymbol{x}^G, \tilde{\boldsymbol{z}})]$), then column 3 represents the percentage difference of the mean of the total cost between ECP-C (SAA) and the greedy policy:

$$(\pi^G - \pi^i)/\pi^G \times 100\%,$$

where $i \in \{E, S\}$. Table 1 also lists the CPU computation time (excluding the evaluation time) of ECP-C and SAA in the last column. The performance of SAA reported here has a sample size of 3,000, as it is the largest sample size where most of the SAA numerical instances can be solved within a reasonable time, between 2 and 10 hours (a larger sample size makes the computation time prohibitively long).

Table 1 Out-of-sample performance of ECP-C and SAA with respect to the greedy policy for different C

C	Method	Total cost		Maximum load				Demand	CPU
		Mean	Std	Overall	On-peak	Mid-peak	Off-peak	charge/π^i	Time (s)
15	ECP-C	12.90%	29.24%	23.33%	23.00%	21.26%	10.13%	0.51	300.42
	SAA	10.97%	21.86%	18.47%	19.51%	15.26%	10.79%	0.53	7,889.92
20	ECP-C	12.12%	25.50%	20.99%	20.68%	18.56%	8.41%	0.53	363.94
	SAA	9.68%	19.95%	16.21%	16.86%	13.19%	5.33%	0.54	7,541.23
25	ECP-C	11.22%	21.80%	18.91%	19.00%	16.03%	8.70%	0.54	453.88
	SAA	8.80%	16.44%	14.46%	15.04%	11.55%	1.80%	0.55	17,740.17
30	ECP-C	10.99%	19.02%	17.94%	18.89%	13.66%	8.59%	0.55	296.55
	SAA	8.55%	14.67%	13.91%	14.44%	11.14%	0.73%	0.56	18,623.61

Notes. The third and fourth columns are the percentage reduction in the mean and standard deviation of the total cost by a given policy compared with the greedy policy, respectively. Analogously, the fifth to eighth columns are the percentage difference of the mean of the maximum load over the entire time horizon, on-peak hours, mid-peak hours, and off-peak hours, respectively; the ninth column is the ratio of the total expected demand charge to the total expected cost.

Value of optimization. As can be seen in Table 1, compared with the greedy policy, ECP-C results in a lower mean and standard deviation of the total cost. For instance, when the number of chargers equals 30, ECP-C results in the mean cost and its standard deviation that are about 10.99% and 19.02% lower than the respective measures under the greedy policy. When station capacity decreases, the difference between ECP-C and the greedy policy is even larger. These differences can be regarded as the value of information on the customer departure time, as ECP-C differs from the greedy policy in that while the former uses this information, the latter does not (recall that the latter charges vehicles with maximal charging power capacity). Our results suggest that a charging station can benefit significantly by incorporating the customer departure information, as our SP model does, to smooth charging electricity load over time and reduce the total cost.

ECP-C versus SAA. Table 1 shows that compared with SAA, ECP-C results in not only a lower expected total cost but also a lower standard deviation of this cost. For instance, when C equals 30, π^E reduces π^G by 10.99%, while π^S reduces π^G by only 8.55%; the standard deviation of π^E reduces that of π^G by 19.02%, while the standard deviation of π^S reduces that of π^G by only 14.67%. This means that the lower cost of π^E is not at the expense of the robustness of ECP-C. This is contrary to the known tradeoff between mean cost and variability: Solutions from a DRO approach usually have a lower variability but a higher mean cost than those from SAA, a

risk-neutral stochastic programming approach (e.g., Choi and Ruszczyński 2008, Hao et al. 2019). ECP-C beats SAA on robustness because ECP-C leverages the ambiguity set inspired by DRO approaches (as noted in §4.3); ECP-C results in a lower mean cost, because while ECP-C is very efficient to compute (using polynomial time solvers), SAA does not scale well in our problem and is hard to solve to optimality even when the sample size is limited (given a sample size of 3000, one out of five numerical instances cannot be solved to the desired accuracy). Compared with π^S, the lower cost π^E can also be seen in the fifth to eighth columns of Table 1: The optimal solution to ECP-C results in lower expected highest load over the entire time horizon and during on-peak, mid-peak, and off-peak hours, and thus ECP-C results in a lower value of the ratio of demand charge to the total expected cost (column 9).

In addition, ECP-C computes much faster than SAA, for instance, about 60 times faster when $C = 30$: While solving SAA takes on average more than 5 hours, ECP-C can be solved in less than 5 minutes. We show the effect of sample size on the performance of SAA in Figure 5 in Appendix B. While SAA is a random approximation, and hence both the computation time and the resulting mean total cost could be sensitive to random samples, ECP-C is a deterministic approximation.

To sum up, compared with SAA, ECP-C not only results in a lower total expected cost but also performs more robustly. In addition, it can be computed much faster. This superiority of ECP-C over SAA continues to hold when we consider other settings—for instance, when the total demand charge consists of only one of the three types, or a mix of these types. Results are omitted for brevity. Therefore, we believe that ECP-C is a better approach than SAA to solve our SP model for the service charging provider.

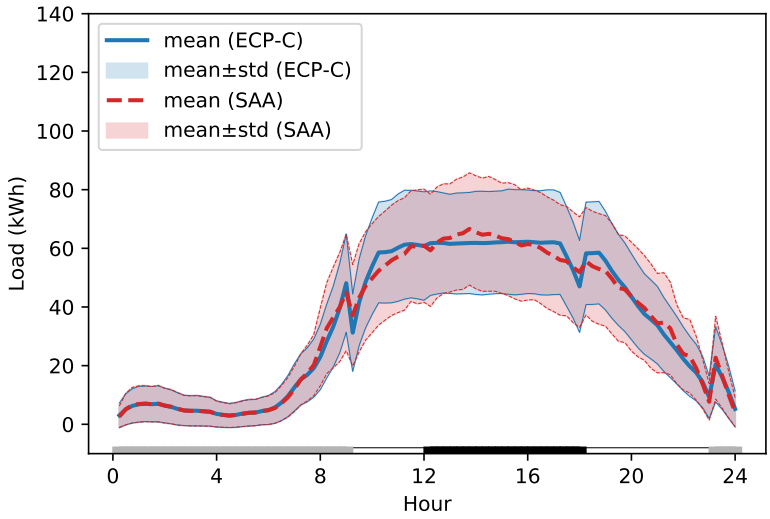
From the perspective of the electric grid, compared with SAA, ECP-C results in a more stable electricity load and lower highest load over time. Figure 3 plots the expected load and its standard deviation over time using ECP-C and SAA given $C = 30$. As can be seen, ECP-C spreads the load more evenly over on-peak and mid-peak hours, and leads to a much lower standard deviation of load during these hours.

Lastly, we estimate the out-of-sample optimality gap of ECP-C (i.e., $(\pi^E - \pi^*)/\pi^* \times 100\%$) using the in-sample π^S as a lower bound of the out-of-sample π^* (Shapiro et al. 2009). For instance, when $C = 30$, this out-of-sample optimality gap of ECP-C is at most $(804.99 - 754.51)/754.51 \times 100\% \approx 6.70\%$, where 804.99 is the out-of-sample π^E and 754.51 is the in-sample π^S.

5.3. Effect of the composition of the total demand charge on electricity load

As discussed above, ECP-C is a better approach to solve our SP model than SAA, and thus we use ECP-C to study the effect of demand charge composition on electricity load over time. Specifically, we examine cases in which the total demand charge is of the following situations: (i) without any

Figure 3 **Mean and standard deviation of electricity load under ECP-C and SAA given** $C = 30$



Notes. The black, white, and gray regions on the x-axis represent on-peak hours (hours 13 to 18), mid-peak hours (hours 9 to 12 and hours 19 to 23), and off-peak hours (hours 1 to 8 and hour 24), respectively.

demand charge; (ii) with only on-peak demand charge; (iii) with only mid-peak demand charge; (iv) with both on-peak and mid-peak demand charges; (v) with only all-period demand charge; and (vi) with both all-period and on-peak demand charges. We compare these with the case in which the total demand charge consists of all three components. Given each composition of the total demand charge, we compute the out-of-sample measures of our SP model evaluated at an optimal solution to the ECP-C approximation, similar to how we evaluate the out-of-sample measures in §5.2. We list the corresponding expected total cost and other performance measures in Table 2 and also plot the mean and standard deviation of electricity load in the charging station for each period in Figure 4. We make the following observations:

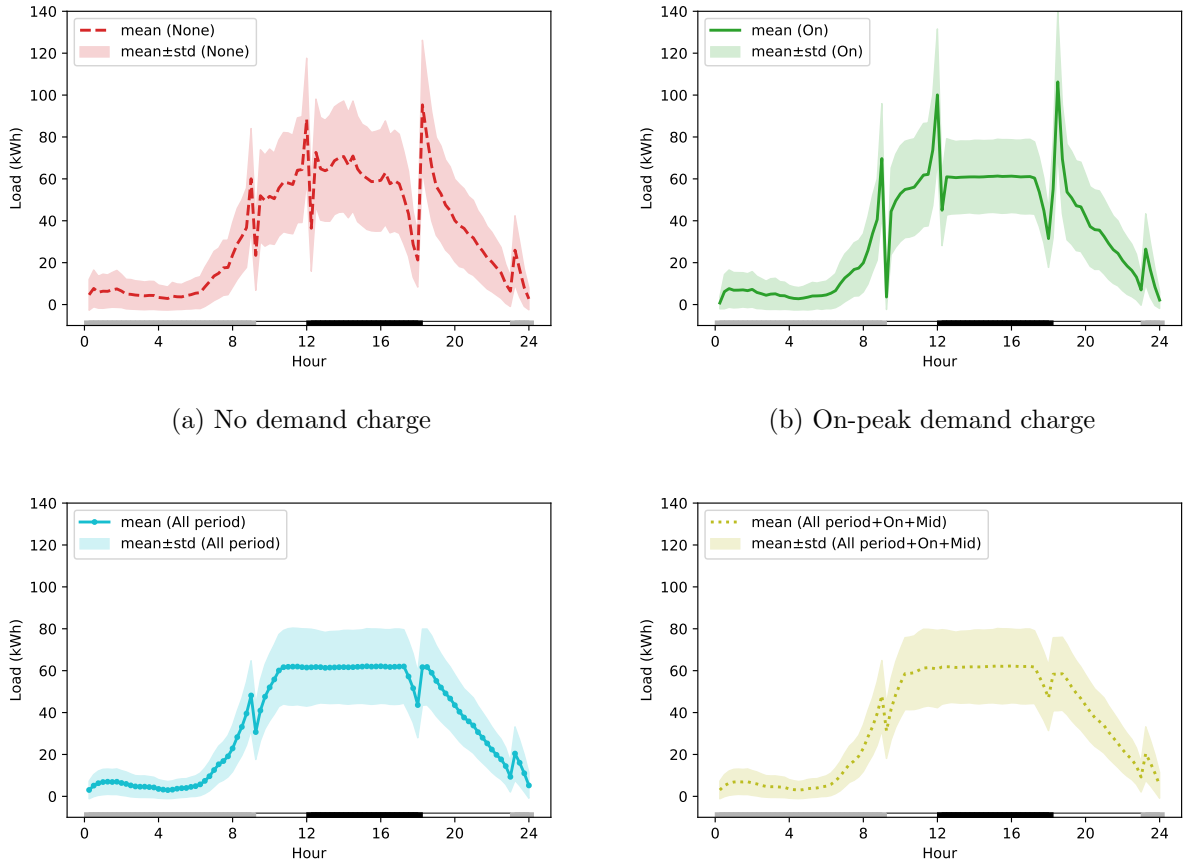
Observation 1. In the absence of any demand charge, the time-of-use energy charge creates an artificial electricity load spike around the transitions between any two time segments. As can be seen in Figure 4a, the expected load swings significantly around the hours between any two sets of time intervals: around hour 9 (between off-peak and mid-peak hours) and around hours 12 and 18 (between mid-peak and on-peak hours). This means that time-of-use energy charge alone is not sufficient to smooth the electricity load and may even create artificial spikes, putting undue stress on the electric grid. This is consistent with the observation in practice (Zhang and Qian 2018): At midnight, there is a huge spike in electricity charging load from residential customers, because their electricity tariff structures do not include any demand charge while their electricity price at midnight is lower than that at hour 11 pm.

Table 2 **Out-of-sample measures of ECP-C for different compositions of the demand charge given** $C = 30$

Tariff	Total cost		Maximum load				Demand
	Mean	Std	Overall	On-peak	Mid-peak	Off-peak	charge/π^E
None	359.25	27.41	127.73	113.85	119.18	66.60	0.00
On	557.79	47.33	128.17	91.92	126.51	73.86	0.35
Mid	418.88	31.32	134.44	133.81	75.42	73.77	0.12
On + Mid	621.99	49.92	101.32	93.00	89.70	73.77	0.42
All-period	545.61	42.51	98.31	93.66	88.23	52.71	0.33
All-period + On	746.97	63.65	98.65	93.15	89.60	53.08	0.51

Notes. "None," "On," "Mid," and "All-period" in the first column represent no, on-peak, mid-peak, and all-period demand charge, respectively.

Figure 4 **Electricity load under ECP-C for different compositions of the demand charge given** $C = 30$



(a) No demand charge

(b) On-peak demand charge

(c) All-period demand charge

(d) All-period, on-peak, and mid-peak demand charge

Observation 2. Either an on-peak or mid-peak demand charge alone cannot smooth the electricity load and may even create a larger load swing than the case without any demand charge. As shown in Figure 4b, when only an on-peak demand charge exists, the electricity load is smoothed only within on-peak hours, but not over time. In particular, there are some artificial spikes in the load around all transition hours (i.e., between on-peak and mid-peak hours and between mid-peak and off-peak hours) because the station simply shifts the high load from on-peak hours to other hours. These spikes are even higher than those within the same hours when there is no demand charge, and thus imposes more stress on the electric grid. When there is only a mid-peak demand charge, an analogous pattern occurs (the corresponding figure is omitted for brevity). Similarly, when both an on-peak demand charge and mid-peak demand charge exist, the load is smoothed only during these combined hours, and load spikes occur around hours between mid-peak and off-peak hours (figure omitted for brevity).

Observation 3. An all-period demand charge alone is sufficient to smooth electricity load over time. In contrast to the case in which there is no demand charge or only either an on-peak or mid-peak demand charge, an all-period demand charge results in a smooth electricity load (see Figure 4c). The load curve for this case almost coincides with that for the case in which all three components of demand charge exist (see Figure 4d). This is because with an all-period demand charge alone, the charging station is incentivized to reduce the highest load over the entire time horizon.

6. Conclusions and future work

We study the EV charging management of a service provider who faces uncertainties in customer arrival/departure times and charging requirements. We formulate this as a two-stage SP with the objective of minimizing the expected total cost in the presence of an electricity tariff that includes demand charges. Due to the large-scale nature of our SP model, SAA, a commom approach to solve such SPs, cannot provide a near-optimal solution within a reasonable amount of time. We then propose using an exponential cone programming approach to solve this SP approximately. When the service provider is uncapacitated, we develop an ECP approximation and show it has an out-of-sample performance guarantee. In the capacitated case, we develop another ECP approximation by also using the idea from DRO to employ entropic dominance ambiguity set and examine its performance on an extensive set of numerical instances calibrated to EV charging data. We find that it is beneficial to model the customer departure information in an SP model like ours, as the ECP and SAA approaches can outperform a greedy charging policy (which ignores such information and is commonly used in practice) by about 11% and 9%, respectively. Moreover, compared with SAA, our ECP approximation results in not only a lower expected total cost but also a lower

standard deviation of this cost. It is also more computationally scalable—ECP runs about sixty times faster given a representative capacity level—and is a deterministic approximation while SAA is sensitive to samples generated. Our results further show the out-of-sample optimality gap of our ECP approximation against the exact SP model is at most 6.7%. Therefore, our ECP approximation is appealing for use in practice by a charging service provider.

On a similar set of data-calibrated numerical instances, we use our ECP approximation to examine how the composition of the total demand charge affects the electricity load over time. We find the following: A time-of-use energy charge alone (i.e., without any demand charge) creates artificial spikes in the electricity load, consistent with the observation in practice. While either an on-peak or mid-peak demand charge alone shifts the artificial spikes and could not smooth electricity load, an all-period demand charge alone can result in load as smooth over time as the case in which all three types of demand charge exist. This insight can give guidance to policy makers in designing the electricity tariff structure.

There are a few promising avenues to extend this work. First, the EV charging schedule in our paper is fixed before uncertainty is realized. In practice, this charging schedule can be adaptive as uncertainty is revealed over time—e.g., by allowing early release of a charger when the customer's charging requirements are fulfilled. A direct extension is to apply our ECP approach in a rolling-horizon fashion. Second, our model does not consider pricing decisions and implicitly assume an exogenously given price for customers to charge their EVs. However, this price could be optimized to maximize the profit for a charging service provider. In this case, our ECP approach could be extended, but SAA does not work any more as it is hard to generate samples when Possion arrival rates depend on the pricing decision. Finally, our methods to construct ECP approximations could also be used to solve other two-stage stochastic linear programs if the epigraph of the MGF of the random variables can be represented using exponential cones or if the MGF of the random variables can be upper bounded by a function whose epigraph can be characterized by exponential cones.

References

Avci B, Girotra K, Netessine S (2014) Electric vehicles with a battery switching station: Adoption and environmental impact. *Management Science* 61(4):772–794.

Ben-Tal A, Den Hertog D, De Waegenaere A, Melenberg B, Rennen G (2013) Robust solutions of optimization problems affected by uncertain probabilities. *Management Science* 59(2):341–357.

Ben-Tal A, El Ghaoui L, Nemirovski A (2009) *Robust Optimization*, volume 28 (Princeton University Press).

Bertsimas D, Sim M (2004) The price of robustness. *Operations Research* 52(1):35–53.

Bertsimas D, Sim M, Zhang M (2019) Adaptive distributionally robust optimization. *Management Science* 65(2):604–618.

Birge JR, Louveaux F (2011) *Introduction to stochastic programming* (Springer Science & Business Media).

Bloomberg (2018) U.S. Electric Vehicle Charging Market to Grow to $18.6 Billion. `https://www.bloomberg.com/news/articles/2018-08-02/u-s-electric-vehicle-charging-market-to-grow-to-18-6-billion`, [Online; accessed 15-Dec-2019].

Boyd S, Vandenberghe L (2004) *Convex Optimization* (Cambridge university press).

Chares R (2009) *Cones and interior-point algorithms for structured convex optimization involving powers and exponentials*. Ph.D. thesis, UCL-Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Chen X, Sim M, Sun P (2007) A robust optimization perspective on stochastic programming. *Operations Research* 55(6):1058–1071.

Chen Z, Sim M, Xu H (2019) Distributionally robust optimization with infinitely constrained ambiguity sets. *Operations Research* 67(5):1328–1344.

Chitkara A, Cross-Call D, Li B, Sherwood J (2016) A review of alternative rate designs. Rocky Mountain Institute, `https://rmi.org/wp-content/uploads/2017/04/A-Review-of-Alternative-Rate-Designs-2016.pdf`, [Online; accessed 15-Dec-2019].

Choi S, Ruszczyński A (2008) A risk-averse newsvendor with law invariant coherent measures of risk. *Operations Research Letters* 36(1):77–82.

Dahl J, Andersen ED (2019) A primal-dual interior-point algorithm for nonsymmetric exponential-cone optimization. Available at Optimization Online.

Danzig GB (1955) Linear programming under uncertainty. *Management Science* 1(3-4):197–206.

Delage E, Ye Y (2010) Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research* 58(3):595–612.

Esfahani PM, Kuhn D (2018) Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming* 171(1-2):115–166.

Goldfarb D, Iyengar G (2003) Robust portfolio selection problems. *Mathematics of operations research* 28(1):1–38.

Hanasusanto GA, Kuhn D, Wiesemann W (2016) A comment on computational complexity of stochastic programming problems. *Mathematical Programming* 159(1-2):557–569.

Hao Z, He L, Hu Z, Jiang J (2019) Robust vehicle pre-allocation with uncertain covariates. *Production and Operations Management* .

He L, Guangrui M, Wei Q, Xin W (2019) Charging electric vehicle sharing fleet. *Manufacturing & Service Operations Management.* Forthcoming.

He L, Mak HY, Rong Y, Shen ZJM (2017) Service region design for urban electric vehicle sharing systems. *Manufacturing & Service Operations Management* 19(2):309–327.

Jaillet P, Loke GG, Sim M (2018) Strategic manpower planning under uncertainty. Available at SSRN 3168168.

Jiang DR, Powell WB (2016) Practicality of nested risk measures for dynamic electric vehicle charging. arXiv preprint arXiv:1605.02848.

Jin C, Tang J, Ghosh P (2013) Optimizing electric vehicle charging with energy storage in the electricity market. *IEEE Transactions on Smart Grid* 4(1):311–320.

JP Morgan (2018) Driving into 2025: The Future of Electric Vehicles. `https://www.jpmorgan.com/global/research/electric-vehicles`, [Online; accessed 15-Dec-2019].

Karlsson S, Kushnir D (2013) How energy efficient is electrified transport. *Systems Perspectives on Electromobility*, chapter 5 (`http://publications.lib.chalmers.se/records/fulltext/179113/local_179113.pdf`), [Online; accessed 17-Dec-2019].

Kleywegt AJ, Shapiro A, Homem-de Mello T (2002) The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization* 12(2):479–502.

Lim MK, Mak HY, Rong Y (2014) Toward mass adoption of electric vehicles: impact of the range and resale anxieties. *Manufacturing & Service Operations Management* 17(1):101–119.

Mak HY, Rong Y, Shen ZJM (2013) Infrastructure planning for electric vehicles with battery swapping. *Management Science* 59(7):1557–1575.

McKinsey & Company (2018) Charging ahead: Electric-vehicle infrastructure demand. `https://perma.cc/9P7A-SP8Q`, [Online; accessed 27-Jan-2020].

MOSEK ApS (2019) *MOSEK Fusion API for Python 9.0.* `https://docs.mosek.com/9.0/pythonfusion/index.html`, [Online; accessed 15-Dec-2019].

National Renewable Energy Laborgy (2017) Identifying Potential Markets for Behind-the-Meter Battery Energy Storage: A Survey of U.S. Demand Charges. `https://www.nrel.gov/docs/fy17osti/68963.pdf`, [NREL/BR-6A20-68963. Online; accessed 27-Jan-2020].

Neubauer J, Simpson M (2015) Deployment of behind-the-meter energy storage for demand charge reduction. Technical report, National Renewable Energy Laboratory (NREL), `https://www.nrel.gov/docs/fy15osti/63162.pdf`, [Online; accessed 15-Dec-2019].

Nissan (2019) 2019 Nissan LEAF range, charging & battery. `https://www.nissanusa.com/vehicles/electric-cars/leaf/features/range-charging-battery.html`, [Online; accessed 15-Dec-2019].

Schneider F, Thonemann UW, Klabjan D (2017) Optimization of battery charging and purchasing at electric vehicle battery swap stations. *Transportation Science* 52(5):1211–1234.

See CT, Sim M (2010) Robust approximation to multiperiod inventory management. *Operations Research* 58(3):583–594.

Shapiro A, Dentcheva D, Ruszczyński A (2009) *Lectures on stochastic programming: modeling and theory* (SIAM).

Shapiro A, Nemirovski A (2005) On complexity of stochastic programming problems. *Continuous optimization*, 111–146 (Springer).

Sun B, Sun X, Tsang DH, Whitt W (2019) Optimal battery purchasing and charging strategy at electric vehicle battery swap stations. *European Journal of Operational Research* 279(2):524 – 539.

UK Department for Transport (2018) Electric Chargepoint Analysis 2017: Local Authority Rapids (revised). `https://www.gov.uk/government/statistics/electric-chargepoint-analysis-2017-local-authority-rapids`, [Online; accessed 15-Dec-2019].

Wiesemann W, Kuhn D, Sim M (2014) Distributionally robust convex optimization. *Operations Research* 62(6):1358–1376.

Wood E, Rames C, Muratori M, Raghavan S, Melaina M (2017) National plug-in electric vehicle infrastructure analysis. National Renewable Energy Laboratory, U.S. Department of Energy, DOE/GO-102017-5040 `https://www.nrel.gov/docs/fy17osti/69031.pdf`, [Online; accessed 29-Jan-2020].

Zap-Map (2019) EV connector types. `https://www.zap-map.com/charge-points/connectors-speeds/`, [Online; accessed 15-Dec-2019].

Zhang P, Qian ZS (2018) User-centric interdependent urban systems: Using time-of-day electricity usage data to predict morning roadway congestion. *Transportation Research Part C: Emerging Technologies* 92:392–411.

Zhang T, Chen W, Han Z, Cao Z (2014) Charging scheduling of electric vehicles with local renewable energy under uncertain electric vehicle arrival and grid power price. *IEEE Transactions on Vehicular Technology* 63(6):2600–2612.

Zhang Y, Lu M, Shen S (2019) On the values of vehicle-to-grid electricity selling in electric vehicle sharing. *Manufacturing & Service Operations Management.* Forthcoming.

Zhu T, Xie J, Sim M (2019) Joint estimation and robustness optimization. Available at SSRN 3335889.

Appendix A: Proofs.

Proof of Theorem 1. For simplicity, we denote $h(\boldsymbol{x}) \triangleq \mathbb{E}_{\mathbb{P}^\infty}[c(\boldsymbol{x}, \tilde{\boldsymbol{z}})]$; for any given $\mu > 0$,

$$\bar{h}(\boldsymbol{x};\mu) \triangleq \sum_{s\in[T]} e_s f_s(\boldsymbol{x},\boldsymbol{\lambda}) + d\left(\mu \ln \sum_{t\in[T]} \exp\left(\sum_{v\in\mathcal{V}_t} \lambda_v \phi\left(x_{v,t}/\mu\right)\right) + \max_{t\in[T]} f_t(\boldsymbol{x},\boldsymbol{\lambda})\right);$$

and

$$\underline{h}(\boldsymbol{x}) \triangleq \sum_{s\in[T]} e_s f_s(\boldsymbol{x},\boldsymbol{\lambda}) + d\left(\max_{t\in[T]} f_t(\boldsymbol{x},\boldsymbol{\lambda})\right).$$

Let $(\bar{\boldsymbol{x}}, \bar{\mu})$ be the optimal solution to $\inf_{\boldsymbol{x}\in\mathcal{X},\mu>0} \bar{h}(\boldsymbol{x};\mu)$, or equivalently the optimal solution to ECP-U by (7) in the proof of Proposition 1, and μ^* be the optimal solution to $\inf_{\mu>0} \bar{h}(\boldsymbol{x}^*;\mu)$. Note that $\underline{h}(\boldsymbol{x})$ is a lower bound of $h(\boldsymbol{x})$ for any $\boldsymbol{x}\in\mathcal{X}$ because $\max_{t\in[T]} f_t(\boldsymbol{x},\boldsymbol{z})$ is convex in $\boldsymbol{z}$. So due to Jensen's inequality, we have

$$\mathbb{E}_{\mathbb{P}^\infty}\left[\max_{t\in[T]} f_t(\boldsymbol{x},\tilde{\boldsymbol{z}})\right] \geq \max_{t\in[T]} f_t(\boldsymbol{x}, \mathbb{E}_{\mathbb{P}^\infty}[\tilde{\boldsymbol{z}}]) = \max_{t\in[T]} f_t(\boldsymbol{x},\boldsymbol{\lambda}).$$

And $\bar{h}(\boldsymbol{x};\mu)$ is an upper bound of $h(\boldsymbol{x})$ from Proposition 1 for any $\mu>0$. Based on the lower bound and upper bound of $\bar{\boldsymbol{x}}$, we bound the performance of $\bar{\boldsymbol{x}}$ in the following.

$$\begin{aligned} h(\bar{\boldsymbol{x}})-h(\boldsymbol{x}^*) &= \big(h(\bar{\boldsymbol{x}})-\bar{h}(\bar{\boldsymbol{x}};\bar{\mu})\big)+\big(\bar{h}(\bar{\boldsymbol{x}};\bar{\mu})-\bar{h}(\boldsymbol{x}^*;\mu^*)\big)+\big(\bar{h}(\boldsymbol{x}^*;\mu^*)-h(\boldsymbol{x}^*)\big)\\ &\le \bar{h}(\boldsymbol{x}^*;\mu^*)-h(\boldsymbol{x}^*)\\ &\le \bar{h}(\boldsymbol{x}^*;\mu^*)-\underline{h}(\boldsymbol{x}^*)\\ &= d\mu^*\ln\sum_{t\in[T]}\exp\left(\sum_{v\in\mathcal{V}_t}\lambda_v\phi(x^*_{v,t}/\mu^*)\right)\\ &= d\inf_{\mu>0}\left(\mu\ln\sum_{t\in[T]}\exp\left(\sum_{v\in\mathcal{V}_t}\lambda_v\phi(x^*_{v,t}/\mu)\right)\right) \end{aligned} \tag{20}$$

where in the first inequality we have $h(\bar{\boldsymbol{x}})\le\bar{h}(\bar{\boldsymbol{x}};\bar{\mu})$ and $\bar{h}(\bar{\boldsymbol{x}};\bar{\mu})\le\bar{h}(\boldsymbol{x}^*;\mu^*)$ due to definition of $(\bar{\boldsymbol{x}},\bar{\mu})$ and $\bar{h}(\boldsymbol{x};\mu)$. Note that for any $\boldsymbol{x}\in\mathcal{X}$ and $\mu>0$, we have

$$\begin{aligned} \sum_{t\in[T]}\exp\left(\sum_{v\in\mathcal{V}_t}\lambda_v\phi(x_{v,t}/\mu)\right) &\le \exp\left(\sum_{t\in[T]}\sum_{v\in\mathcal{V}_t}\lambda_v\phi(x_{v,t}/\mu)\right)+T\\ &= \exp\left(\sum_{v\in[V]}\lambda_v\sum_{t\in\mathcal{T}_v}\phi(x_{v,t}/\mu)\right)+T\\ &\le \exp\left(\sum_{v\in[V]}\lambda_v\phi\left(\sum_{t\in\mathcal{T}_v}x_{v,t}/\mu\right)\right)+T\\ &= \exp\left(\sum_{v\in[V]}\lambda_v\phi\left(\frac{u_v}{\eta\mu}\right)\right)+T \end{aligned} \tag{21}$$

where the first inequality is due to $e^{x_1}+e^{x_2}\le e^{x_1+x_2}+1$ for any $x_1,x_2\ge0$ and the second inequality is from the super-additivity of $\phi(x)$ for $x\ge0$, i.e., $\phi(x_1)+\phi(x_2)\le\phi(x_1+x_2),\forall x_1,x_2\ge0$. Therefore, by combining the bounds (20) and (21), we get for any $\mu>0$,

$$\begin{aligned} h(\bar{\boldsymbol{x}})-h(\boldsymbol{x}^*) &\le d\mu\ln\left(\exp\left(\sum_{v\in[V]}\lambda_v\phi\left(\frac{u_v}{\eta\mu}\right)\right)+T\right)\\ &\le d\mu\ln\left(2T\exp\left(\sum_{v\in[V]}\lambda_v\phi\left(\frac{u_v}{\eta\mu}\right)\right)\right)\\ &\le d\mu\left(\sum_{v\in[V]}\lambda_v\phi\left(\frac{u_v}{\eta\mu}\right)+\ln(2T)\right), \end{aligned} \tag{22}$$

where the second inequality is from the fact that $\ln(x_1+x_2)\le\ln(x_1x_2+1)\le\ln(2x_1x_2)$ for any $x_1,x_2\ge1$. So by taking the minimum over μ, we obtain the performance guarantee for $h(\bar{\boldsymbol{x}})$ in (13).

Since the inequality (22) holds for all $\mu>0$, so by choosing

$$\mu:=\frac{1}{\eta}\sqrt{\sum_{v\in[V]}\frac{\lambda_v u_v^2}{2\ln(2T)}},$$

the bound (22) becomes

$$h(\bar{\boldsymbol{x}})-h(\boldsymbol{x}^*)\le\frac{d}{\eta}\sqrt{\sum_{v\in[V]}\frac{\lambda_v u_v^2}{2\ln(2T)}}\sum_{v\in[V]}\lambda_v\phi\left(\frac{u_v\sqrt{2\ln(2T)}}{\sqrt{\sum_{v\in[V]}\lambda_v u_v^2}}\right)+\frac{d\ln(2T)}{\eta}\sqrt{\sum_{v\in[V]}\frac{\lambda_v u_v^2}{2\ln(2T)}}. \tag{23}$$

We can then estimate the above expression by approximating

$$\phi\left(\frac{u_v\sqrt{2\ln(2T)}}{\sqrt{\sum_{v\in[V]}\lambda_v u_v^2}}\right) \le \frac{2u_v^2\ln(2T)}{\sum_{v\in[V]}\lambda_v u_v^2},$$

when $(2u_v^2\ln(2T))/\sum_{v\in[V]}\lambda_v u_v^2 \le 1$ for all $v\in[V]$ as $\phi(x)\le x^2$ for any $x\le 1$. Then the bound (23) can be bounded by

$$\left(\sqrt{2}+\frac{1}{\sqrt{2}}\right)\frac{d}{\eta}\sqrt{\ln(2T)\sum_{v\in[V]}\lambda_v u_v^2}.$$

□

Proof of Lemma 1:

$\mathcal{F}^1\cap\mathcal{F}^2\subseteq\mathcal{F}$ is clearly true because $\mathcal{F}^1\cap\mathcal{F}^2$ includes all the constraints in $\mathcal{F}$. To prove $\mathcal{F}\subseteq\mathcal{F}^1\cap\mathcal{F}^2$, we next show that any $\tilde{\boldsymbol{z}}$ that satisfies the following also satisfies $\mathbb{E}_{\mathbb{P}}[\tilde{\boldsymbol{z}}]\le\boldsymbol{\lambda}$:

$$\ln\mathbb{E}_{\mathbb{P}}\left[\exp\left(\boldsymbol{\theta}'\tilde{\boldsymbol{z}}\right)\right] \le \sum_{v\in[V]}\lambda_v\left(e^{\theta_v}-1\right)\ \forall\boldsymbol{\theta}\ge\mathbf{0}$$

For each $v\in[V]$, consider $\boldsymbol{\theta}_v\triangleq(0,...,\theta_v,0,...,0)$, i.e., every component is 0 except one, which is strictly positive. Then we have $\ln\mathbb{E}_{\mathbb{P}}\left[\exp\left(\boldsymbol{\theta}'\tilde{\boldsymbol{z}}\right)\right]=\ln\mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta_v\tilde{z}_v\right)\right]\le\lambda_v(e^{\theta_v}-1)$. Taking exponential function on both sides we have an equivalent inequality

$$\mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta_v\tilde{z}_v\right)\right]\le\exp\left(\lambda_v(e^{\theta_v}-1)\right).$$

Using Taylor's expansion on both sides of the inequality, we have for each $\theta_v>0$

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}\left[\exp\left(\theta_v\tilde{z}_v\right)\right] &= \mathbb{E}_{\mathbb{P}}\left[1+\theta_v\tilde{z}_v+\sum_{k=2}^{\infty}\frac{(\theta_v\tilde{z}_v)^k}{k!}\right]\\
&= 1+\mathbb{E}_{\mathbb{P}}\left[\theta_v\tilde{z}_v\right]+\sum_{k=2}^{\infty}\mathbb{E}_{\mathbb{P}}\left[\frac{(\theta_v\tilde{z}_v)^k}{k!}\right]\\
&= 1+\mathbb{E}_{\mathbb{P}}\left[\theta_v\tilde{z}_v\right]+o(\theta_v)\\
&\le \exp\left(\lambda_v(e^{\theta_v}-1)\right)\\
&= 1+\lambda_v\theta_v+o(\theta_v),
\end{aligned}$$

where the second equality is from Fubini's theorem. After simplifying and dividing θ_v on both sides of

$$\mathbb{E}_{\mathbb{P}}\left[\theta_v\tilde{z}_v\right]+o(\theta_v)\le\lambda_v\theta_v+o(\theta_v),$$

then letting $\theta_v\to 0^+$, we get $\mathbb{E}_{\mathbb{P}}\left[\tilde{z}_v\right]\le\lambda_v$ for each $v\in[V]$. Hence $\mathbb{E}_{\mathbb{P}}\left[\tilde{\boldsymbol{z}}\right]\le\boldsymbol{\lambda}$. □

To prove Propostion 3, we establish the next lemma first.

LEMMA 2. *For a piece-wise affine convex function* $g(\boldsymbol{z})\triangleq\max_{i\in[I]}\left\{\bar{a}_i+\boldsymbol{a}_i'\boldsymbol{z}\right\}$ *and any random vector* $\tilde{\boldsymbol{z}}\sim\mathbb{P}$, *we have*

$$\mathbb{E}_{\mathbb{P}}\left[g(\tilde{\boldsymbol{z}})\right]=\inf_{\bar{b}_i,\boldsymbol{b}_i,i\in[I]}\mathbb{E}_{\mathbb{P}}\left[\max_{i\in[I]}\left\{\bar{b}_i+\boldsymbol{b}_i'\tilde{\boldsymbol{z}}\right\}\right]+\mathbb{E}_{\mathbb{P}}\left[\max_{i\in[I]}\left\{(\bar{a}_i-\bar{b}_i)+(\boldsymbol{a}_i-\boldsymbol{b}_i)'\tilde{\boldsymbol{z}}\right\}\right]. \tag{24}$$

Proof: Note (24) holds when $\bar{b}_i = 0$ and $\boldsymbol{b}_i = \mathbf{0}$ for all $i \in [I]$, we have

$$\mathbb{E}_{\mathbb{P}}\left[g(\tilde{\boldsymbol{z}})\right] \geq \inf_{\bar{b}_i, \boldsymbol{b}_i, i\in[I]} \mathbb{E}_{\mathbb{P}}\left[\max_{i\in[I]}\left\{\bar{b}_i + \boldsymbol{b}_i'\tilde{\boldsymbol{z}}\right\}\right] + \mathbb{E}_{\mathbb{P}}\left[\max_{i\in[I]}\left\{(\bar{a}_i - \bar{b}_i) + (\boldsymbol{a}_i - \boldsymbol{b}_i)'\tilde{\boldsymbol{z}}\right\}\right].$$

Note that for all $\boldsymbol{z}$ and $\bar{b}_i, \boldsymbol{b}_i, i \in [I]$, we have

$$g(\boldsymbol{z}) \leq \max_{i\in[I]}\left\{\bar{b}_i + \boldsymbol{b}_i'\boldsymbol{z}\right\} + \max_{i\in[I]}\left\{(\bar{a}_i - \bar{b}_i) + (\boldsymbol{a}_i - \boldsymbol{b}_i)'\boldsymbol{z}\right\}.$$

Hence for all $\bar{b}_i$ and $\boldsymbol{b}_i, i \in [I]$, we have

$$\mathbb{E}_{\mathbb{P}}\left[g(\tilde{\boldsymbol{z}})\right] \leq \mathbb{E}_{\mathbb{P}}\left[\max_{i\in[I]}\left\{\bar{b}_i + \boldsymbol{b}_i'\tilde{\boldsymbol{z}}\right\}\right] + \mathbb{E}_{\mathbb{P}}\left[\max_{i\in[I]}\left\{(\bar{a}_i - \bar{b}_i) + (\boldsymbol{a}_i - \boldsymbol{b}_i)'\tilde{\boldsymbol{z}}\right\}\right].$$

By taking infimum over $\bar{b}_i$ and $\boldsymbol{b}_i, i \in [I]$, we get the conclusion. □

Proof of Proposition 3

Since DRO-Ent is an upper bound of (4), we then obtain an upper bound of DRO-Ent, which then becomes an upper bound of (4). We can then bound $\sup_{\mathbb{P}\in\mathcal{F}} \mathbb{E}_{\mathbb{P}}\left[c(\boldsymbol{x}, \tilde{\boldsymbol{z}})\right]$ as follow,

$$\sup_{\mathbb{P}\in\mathcal{F}} \mathbb{E}_{\mathbb{P}}\left[c(\boldsymbol{x}, \tilde{\boldsymbol{z}})\right] \leq \sup_{\mathbb{P}\in\mathcal{F}} \mathbb{E}_{\mathbb{P}}\left[\sum_{s\in[T]} e_s f_s(\boldsymbol{x}, \tilde{\boldsymbol{z}})\right] + d \sup_{\mathbb{P}\in\mathcal{F}} \mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}\{f_t(\boldsymbol{x}, \tilde{\boldsymbol{z}})\}\right]. \tag{25}$$

Then we derive an upper bound for each of the two terms on the RHS of (25). For the first term, due to the fact $\mathcal{F} \subseteq \mathcal{F}^2$ (based on Lemma 1), we have

$$\sup_{\mathbb{P}\in\mathcal{F}} \mathbb{E}_{\mathbb{P}}\left[\sum_{s\in[T]} e_s f_s(\boldsymbol{x}, \tilde{\boldsymbol{z}})\right] \leq \sup_{\mathbb{P}\in\mathcal{F}^2} \mathbb{E}_{\mathbb{P}}\left[\sum_{s\in[T]} e_s f_s(\boldsymbol{x}, \tilde{\boldsymbol{z}})\right].$$

Using standard duality result, we have

$$\begin{aligned}
\sup_{\mathbb{P}\in\mathcal{F}^2} \mathbb{E}_{\mathbb{P}}\left[\sum_{s\in[T]} e_s f_s(\boldsymbol{x}, \tilde{\boldsymbol{z}})\right] &= \inf_{\boldsymbol{b}\geq\mathbf{0}} \; a + \boldsymbol{b}'\boldsymbol{\lambda} \\
&\quad \text{s.t.} \sum_{s\in[T]} e_s f_s(\boldsymbol{x}, \boldsymbol{z}) \leq a + \boldsymbol{b}'\boldsymbol{z} \qquad \forall \boldsymbol{z} \in \mathcal{Z} \\
&= \inf_{\boldsymbol{b}\geq\mathbf{0}} \; a + \boldsymbol{b}'\boldsymbol{\lambda} \\
&\quad \text{s.t.} \sup_{\boldsymbol{z}\in\mathcal{Z}} \sum_{s\in[T]} e_s f_s(\boldsymbol{x}, \boldsymbol{z}) - \boldsymbol{b}'\boldsymbol{z} \leq a
\end{aligned} \tag{26}$$

Note that the dual of $\sup_{\boldsymbol{z}\geq\mathbf{0}}\left\{\sum_{s\in[T]} e_s \sum_{v\in\mathcal{V}_s} x_{v,s} z_v - \sum_{v\in[V]} b_v z_v \mid \sum_{v\in\mathcal{V}_t} z_v \leq C, \forall t \in [T]\right\} \leq a$ is

$$\inf_{\boldsymbol{\nu}\geq\mathbf{0}}\left\{C\sum_{t\in[T]}\nu_t \mid \sum_{s\in\mathcal{T}_v} x_{v,s}e_s - b_v \leq \sum_{t\in\mathcal{T}_v}\nu_t, \forall v \in [V]\right\} \leq a,$$

which is equivalent to

$$\begin{aligned}
&C\textstyle\sum_{t\in[T]}\nu_t \leq a \\
&\textstyle\sum_{s\in\mathcal{T}_v} x_{v,s}e_s - b_v \leq \sum_{t\in\mathcal{T}_v}\nu_t \;\; \forall v \in [V].
\end{aligned}$$

Therefore, (26) becomes

$$\begin{aligned}\sup_{\mathbb{P}\in\mathcal{F}^2}\mathbb{E}_{\mathbb{P}}\left[\sum_{s\in[T]}e_s f_s(\boldsymbol{x},\tilde{\boldsymbol{z}})\right] = \inf_{\boldsymbol{b}\geq\boldsymbol{0},\boldsymbol{\nu}\geq\boldsymbol{0}}\ & a+\boldsymbol{b}'\boldsymbol{\lambda}\\ \text{s.t.}\quad & C\sum_{t\in[T]}\nu_t\leq a\\ & \sum_{s\in\mathcal{T}_v}x_{v,s}e_s-b_v\leq\sum_{t\in\mathcal{T}_v}\nu_t\ \ \forall v\in[V]\end{aligned}\tag{27}$$

We next obtain an upper bound of $\sup_{\mathbb{P}\in\mathcal{F}}\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}\{f_t(\boldsymbol{x},\tilde{\boldsymbol{z}})\}\right]$. Since $c(\boldsymbol{x},\boldsymbol{z})$ is a piece-wise linear convex function of $\boldsymbol{z}$ given $\boldsymbol{x}$, we have

$$\begin{aligned}\sup_{\mathbb{P}\in\mathcal{F}}\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}\{f_t(\boldsymbol{x},\tilde{\boldsymbol{z}})\}\right] &= \sup_{\mathbb{P}\in\mathcal{F}}\inf_{\boldsymbol{y}}\left\{\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}\{f_t(\boldsymbol{y},\tilde{\boldsymbol{z}})\}\right]+\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}\{f_t(\boldsymbol{x}-\boldsymbol{y},\tilde{\boldsymbol{z}})\}\right]\right\}\\ &\leq \inf_{\boldsymbol{y}}\sup_{\mathbb{P}\in\mathcal{F}}\left\{\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}\{f_t(\boldsymbol{y},\tilde{\boldsymbol{z}})\}\right]+\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}\{f_t(\boldsymbol{x}-\boldsymbol{y},\tilde{\boldsymbol{z}})\}\right]\right\}\\ &\leq \inf_{\boldsymbol{y}}\left\{\sup_{\mathbb{P}\in\mathcal{F}^1}\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}\{f_t(\boldsymbol{y},\tilde{\boldsymbol{z}})\}\right]+\sup_{\mathbb{P}\in\mathcal{F}^2}\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}\{f_t(\boldsymbol{x}-\boldsymbol{y},\tilde{\boldsymbol{z}})\}\right]\right\}\\ &\leq \inf_{\boldsymbol{y}\geq\boldsymbol{0}}\left\{\sup_{\mathbb{P}\in\mathcal{F}^1}\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}\{f_t(\boldsymbol{y},\tilde{\boldsymbol{z}})\}\right]+\sup_{\mathbb{P}\in\mathcal{F}^2}\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}\{f_t(\boldsymbol{x}-\boldsymbol{y},\tilde{\boldsymbol{z}})\}\right]\right\},\end{aligned}\tag{28}$$

where the equality is from Lemma 2 and the second inequality is from Lemma 1. We next bound $\sup_{\mathbb{P}\in\mathcal{F}^1}\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}f_t(\boldsymbol{x},\tilde{\boldsymbol{z}})\right]$ and $\sup_{\mathbb{P}\in\mathcal{F}^2}\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}f_t(\boldsymbol{x},\tilde{\boldsymbol{z}})\right]$. Similar to the proof of Proposition 1, for any $\boldsymbol{y}\geq\boldsymbol{0}$, we have

$$\begin{aligned}\sup_{\mathbb{P}\in\mathcal{F}^1}\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}f_t(\boldsymbol{y},\tilde{\boldsymbol{z}})\right] &= \sup_{\mathbb{P}\in\mathcal{F}^1}\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}\{f_t(\boldsymbol{y},\tilde{\boldsymbol{z}})-f_t(\boldsymbol{y},\boldsymbol{\lambda})+f_t(\boldsymbol{y},\boldsymbol{\lambda})\}\right]\\ &\leq \sup_{\mathbb{P}\in\mathcal{F}^1}\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}\{f_t(\boldsymbol{y},\tilde{\boldsymbol{z}})-f_t(\boldsymbol{y},\boldsymbol{\lambda})\}\right]+\max_{t\in[T]}f_t(\boldsymbol{y},\boldsymbol{\lambda})\\ &\leq \mu\ln\sum_{t\in[T]}\sup_{\mathbb{P}\in\mathcal{F}^1}\mathbb{E}_{\mathbb{P}}\left[\exp\left((f_t(\boldsymbol{y},\tilde{\boldsymbol{z}})-f_t(\boldsymbol{y},\boldsymbol{\lambda}))/\mu\right)\right]+\max_{t\in[T]}f_t(\boldsymbol{y},\boldsymbol{\lambda})\\ &\leq \mu\ln\sum_{t\in[T]}\exp\left(\sum_{v\in\mathcal{V}_t}\lambda_v\left(e^{y_{v,t}/\mu}-1-y_{v,t}/\mu\right)\right)+\max_{t\in[T]}f_t(\boldsymbol{y},\boldsymbol{\lambda}),\end{aligned}$$

which has the same form as (8). By the same argument as in the proof of Proposition 1, we can bound $\sup_{\mathbb{P}\in\mathcal{F}^1}\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}f_t(\boldsymbol{y},\tilde{\boldsymbol{z}})\right]$ by

$$\begin{aligned}\inf_{\kappa,\gamma,\mu>0,\boldsymbol{\xi},\boldsymbol{\zeta}}\ & \sum_{s\in[T]}e_s f_s(\boldsymbol{y},\boldsymbol{\lambda})+d\left(\kappa+\gamma\right)\\ \text{s.t.}\quad & \sum_{v\in\mathcal{V}_t}y_{v,t}\lambda_v\leq\gamma && \forall t\in[T]\\ & \mu\exp\left(y_{v,t}/\mu\right)\leq\xi_{v,t} && \forall t\in[T],v\in\mathcal{V}_t\\ & \mu\exp\left(\left(-\kappa+\sum_{v\in\mathcal{V}_t}\lambda_v(\xi_{v,t}-y_{v,t}-\mu)\right)/\mu\right)\leq\zeta_t && \forall t\in[T]\\ & \sum_{t\in[T]}\zeta_t\leq\mu\end{aligned}\tag{29}$$

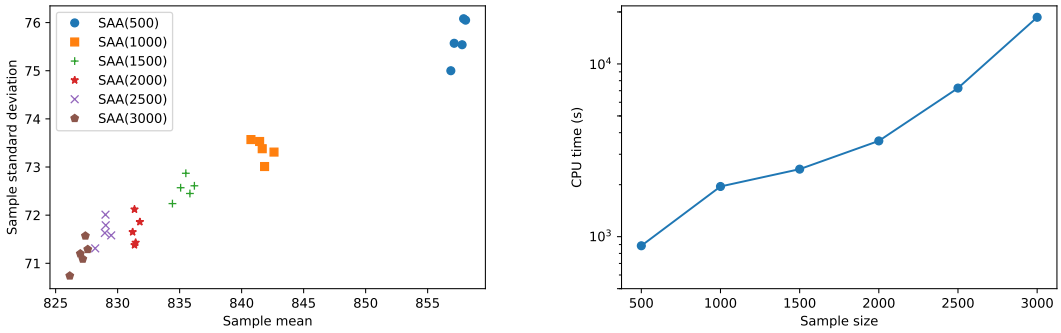
Similar to (27), we bound $\sup_{\mathbb{P}\in\mathcal{F}^2}\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}\{f_t(\boldsymbol{x},\tilde{\boldsymbol{z}})\}\right]$ using the strong duality results, i.e.,

$$\begin{aligned}
\sup_{\mathbb{P}\in\mathcal{F}^2}\mathbb{E}_{\mathbb{P}}\left[\max_{t\in[T]}\{f_t(\boldsymbol{x},\tilde{\boldsymbol{z}})\}\right] = \quad & \inf_{\boldsymbol{\beta}\geq\mathbf{0}} \quad \alpha+\boldsymbol{\beta}'\boldsymbol{\lambda} \\
& \text{s.t.} \quad f_t(\boldsymbol{x},\boldsymbol{z})\leq\alpha+\boldsymbol{\beta}'\boldsymbol{z} \qquad \forall\boldsymbol{z}\in\mathcal{Z}, t\in[T] \\
= \quad & \inf_{\boldsymbol{\beta}\geq\mathbf{0}} \quad \alpha+\boldsymbol{\beta}'\boldsymbol{\lambda} \\
& \text{s.t.} \quad \sup_{\boldsymbol{z}\in\mathcal{Z}} f_t(\boldsymbol{x},\boldsymbol{z})-\boldsymbol{\beta}'\boldsymbol{z}\leq\alpha \ \forall t\in[T] \\
= \quad & \inf_{\boldsymbol{\rho}\geq\mathbf{0},\boldsymbol{\beta}\geq\mathbf{0}} \quad \alpha+\boldsymbol{\beta}'\boldsymbol{\lambda} \\
& \text{s.t.} \quad C\sum_{k\in[T]}\rho_t^k\leq\alpha \qquad \forall t\in[T] \\
& \qquad\quad x_{v,t}-\beta_v\leq\sum_{k\in\mathcal{T}_v}\rho_t^k \qquad \forall t\in[T], v\in\mathcal{V}_t \\
& \qquad\quad -\beta_v\leq\sum_{k\in\mathcal{T}_v}\rho_t^k \qquad \forall t\in[T], v\in[V]\backslash\mathcal{V}_t
\end{aligned} \tag{30}$$

Note $-\beta_v\leq\sum_{k\in\mathcal{T}_v}\rho_t^k$ are redundant. Combining bound (25), (27), (28), (29), and (30) and optimizing over $\boldsymbol{x}$, we get an upper bound of the objective function $\sup_{\mathbb{P}\in\mathcal{F}}\mathbb{E}_{\mathbb{P}}[c(\boldsymbol{x},\tilde{\boldsymbol{z}})]$ for any $\boldsymbol{x}\in\mathcal{X}$ as ECP-C. □

Appendix B: Performance of SAA with different sample sizes

Figure 5 **Performance of SAA with different sample sizes given** $C=30$



(a) SAA mean and std at different sample sizes

(b) CPU time at different sample sizes