Data-Driven Distributionally Robust Electric Vehicle Balancing for Autonomous Mobility-on-Demand Systems under Demand and Supply Uncertainties

Sihong He, Zhili Zhang, Shuo Han, Lynn Pepin, Guang Wang, Desheng Zhang, John Stankovic, Fei Miao

***Abstract*—Electric vehicles (EVs) are being rapidly adopted due to their economic and societal benefits. Autonomous mobility-on-demand (AMoD) systems also embrace this trend. However, the long charging time and high recharging frequency of EVs pose challenges to efficiently managing EV AMoD systems. The complicated dynamic charging and mobility process of EV AMoD systems makes the demand and supply uncertainties significant when designing vehicle balancing algorithms. In this work, we design a data-driven distributionally robust optimization (DRO) approach to balance EVs for both the mobility service and the charging process. The optimization goal is to minimize the worst-case expected cost under both passenger mobility demand uncertainties and EV supply uncertainties. We then propose a novel distributional uncertainty sets construction algorithm that guarantees the produced parameters are contained in desired confidence regions with a given probability. To solve the proposed DRO AMoD EV balancing problem, we derive an equivalent computationally tractable convex optimization problem. Based on real-world EV data of a taxi system, we show that with our solution the average total balancing cost is reduced by 14.49%, and the average mobility fairness and charging fairness are improved by 15.78% and 34.51%, respectively, compared to solutions that do not consider uncertainties.**

***Index Terms*—Data Driven, Electric Vehicle, Mobility-on-Demand Systems, Fairness, Distributionally Robust Optimization**

I. Introduction

In Autonomous Mobility-on-Demand (AMoD) systems, self-driving vehicles provide personal on-demand transportation service for customers and rebalance themselves to maintain acceptable quality of service throughout the system [1], [2]. AMoD systems have been advocated as one of the most promising energy-efficient transportation solutions. Electric Vehicles (EVs) have tremendous potential in AMoD systems for being economical and environmentally friendly [3]. For instance, EV AMoD systems directly address the problems of oil dependency and air pollution. However, EVs have quite different energy-refilling patterns compared with traditional gas-powered vehicles. They have long charging times, high charging frequency, uncertain sporadic demands, and dispersed mobility patterns [4], [5].

There are emerging problems when commercial EVs are gradually introduced into AMoD systems, considering current charging technologies and limited charging infrastructures. EVs' frequent intermittent charging requirements may reduce the quality of mobility service. Unbalanced EV distribution may also cause long waiting times and low charging service quality in some charging stations. To address these challenges, researchers proposed plenty of vehicle allocation methods, charging scheduling approaches and joint charging-relocation recommendation schemes. Vehicle allocation methods rebalance the vehicle distributions over time in responding to known or predicted demand and supply [1], [6], [7]. We further discuss these methods in the Related Work section.

However, most existing vehicle allocation methods and charging scheduling approaches do not consider uncertainties by assuming the measurement and prediction models are perfect [1], [8], while model uncertainty affects the performance of decisions [9]. And it is difficult to accurately predict passengers' demand and EVs' charging patterns. As the promotion of EVs continues, we cannot ignore the uncertainties caused by EVs' charging behaviors. For example, uncertain charging time, queuing time and charging frequency all contribute to the supply uncertainty, due to the limited knowledge we have about charging patterns [4]. Then difficulties appear in many aspects, such as introducing multiple uncertainties into AMoD systems, modeling the uncertain parameters, and analyzing the mutual dependency between supply and demand. Therefore, making real-time decisions under both supply and demand uncertainties is still a charging and unsolved research problem.

Considering both passenger demand and EV supply uncertainties, we propose a distributionally robust optimization (DRO) approach to make robust vehicle balancing decisions for both mobility service and charging scheduling. DRO method considers the uncertain parameters' probability distributions are contained in some pre-specified distributional uncertainty sets [9]. More DRO literature is discussed in Section II. In our method, we assume the true probability distribution of passenger demand and EV supply lies in a set of probability distributions, i.e., a distributional uncertainty set. We define the vehicle balancing cost and system-level charging service fairness requirement in our objective function, which is convex over the decision variables and concave over the uncertain

This work is partially supported by NSF S&AS-1849246, NSF CPS-1932250 and NSF S&CC-1952096. Sihong He, Zhili Zhang, Lynn Pepin, and Fei Miao are with the Department of Computer Science and Engineering, University of Connecticut, Storrs Mansfield, CT, USA 06268. Email: {sihong.he, zhili.zhang, lynn.pepin, fei.miao}@uconn.edu. This work is also partially supported by NSF S&AS-1849238 and CPS-1932223. Shuo Han is with the Department of Electrical and Computer Engineering, University of Illinois, Chicago, IL, USA 60607. Email:hanshuo@uic.edu. Guang Wang is with the Department of Computer Science, Florida State University, Tallahassee, FL 32304. Email: guang@cs.fsu.edu. Desheng Zhang is with the Department of Computer Science, Rutgers University, Piscataway, NJ, USA 08901. Email: desheng.zhang@cs.rutgers.edu. John A. Stankovic is with the University of Virginia, Charlottesville, VA, USA 22904. Email: stankovic@cs.virginia.edu.

parameters. The vehicle balancing cost includes the balancing cost to send vacant EVs to predicted mobility demand and low-battery EVs to charging stations. We put the mobility service fairness requirement in convex constraints. This objective and constraint design decouple the mutual dependencies between EV supply and passenger demand. Thus, we calculate balancing decisions by solving the DRO problem, i.e. minimizing the worst-case expected objective function over distributional uncertainty sets and convex constraints. We further derive an equivalent convex optimization problem form for the DRO problem to provide solutions in a computationally tractable way. We also propose an efficient distributional uncertainty set construction algorithm to construct stable uncertainty sets. We briefly summarize the proposed framework structure of this paper in Fig. 1.

The *key contributions* of our work are as follows:

- To the best of our knowledge, our proposed mathematical system-level vehicle balancing framework is the first to consider both future mobility demand uncertainties and EV supply uncertainties for EV AMoD systems. While model predictive control algorithms [1], [10], [11] have been designed considering AMoD system demand uncertainties in the literature, the supply side uncertainties for EV AMoD are not well studied yet.
- We design a distributionally robust optimization approach to balance EVs across a city to provide fair passenger mobility and EV charging service while reducing the total balancing cost. We consider probabilistic distribution uncertainties of both the passenger mobility demand and the EV supply caused by the challenge of charging process prediction [5], [12]. The proposed problem formulation decouples the mutual dependencies between EV supply and passenger demand. We further design an efficient algorithm to construct distributional uncertainty sets, and prove that the produced uncertainty set parameters are guaranteed being contained in desired confidence regions with a given probability.
- We derive an equivalent convex optimization problem for the proposed distributionally robust optimization problem. Hence, we provide a system-level performance guarantee in a computationally tractable way under supply and demand uncertainties. Based on the E-taxi fleet of Shenzhen city, which is a real-world EV AMoD system dataset, we show that our method reduces the average total balancing cost, the average mobility unfairness and charging unfairness by 14.49%, 15.78% and 34.51%, respectively, compared to non-robust solutions.

The rest of the paper is organized as follows. The related work and distributionally robust EV balancing problem formulation are presented in Section II and III, respectively. The formal distribuitonally uncertainty set form and novel construction algorithms are in Section IV. An equivalent computationally tractable form is derived in Section V. Experiments are in Section VI. We conclude in Section VII.

II. Related Work

A. Vehicle Allocation: To improve the performance of AMoD systems, multiple vehicle allocation and balancing approaches have been proposed. For instance, queuing network model [8], flow framework [13], model predictive control [1], [10], receding horizon control [14], and reinforcement learning method [15], [16] have been designed. However, most of them do not consider EV charging patterns nor uncertainties caused by EV charging behaviors. Making real-time decisions under supply and demand uncertainties is still a challenging and unsolved problem. Our work jointly considers the EV allocation and charging problem in a distributionally robust optimization problem while considering two-sided uncertainties.

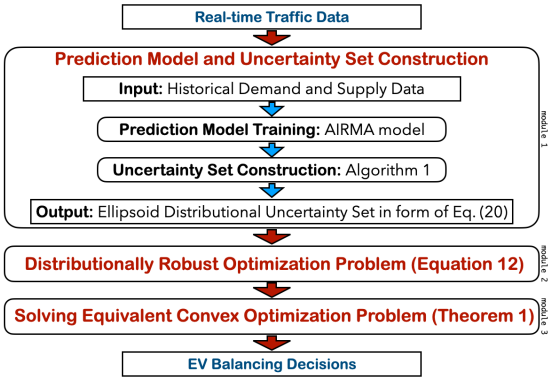


Fig. 1. We briefly summarize the proposed framework structure of this paper in this figure. In module 1, we use historical data to train a prediction model for supply and demand then use the well-trained prediction models as the input of our Algorithm 1 to finally get the distributional uncertainty sets for demand and supply in form of (20). In module 2, we get the DRO EV balancing problem (the distributionally robust optimization problem (12)) using the real-time sensing data and distributional uncertainty sets. In module 3, we apply the conclusions in Theorem 1 to obtain the EV balancing decisions by solving the equivalent convex optimization problem (25).

B. Charging Scheduling: To improve EV charging efficiency, MDP based [17] and queuing model based [5] charging scheduling, charging station deployment [18], online charging recommendations [12] have been proposed. Future charging supply or demand are usually considered in charging recommendation [19], [20]. But these methods haven't provided integrated passenger picking-up and EV charging scheduling solutions for EV AMoD systems. Though an E-taxi charging framework under dynamics of renewable energy and passenger mobility is proposed [19], [21], its performance may be undermined by system uncertainties that are not considered. Other existing works either only focus on robust charging scheduling with uncertainties from charging behavior [22], [23] or only study joint EV balancing and charging strategies without considering uncertainties [19]. It is still challenging to simultaneously deal with passenger mobility demand and EV supply uncertainties when making EV balancing decisions for both passenger picking-up and EV charging. Our work fills this gap by proposing a novel distributionally robust optimization EV balancing framework to provide fair passenger mobility and EV charging service while reducing the total balancing cost.

C. Robust and Distributionally Robust Optimization: Robust optimization (RO) assumes that uncertain parameters can be any value in an uncertainty set, whereas distributionally robust optimization (DRO) models the uncertain parameters as random variables whose underlying probability distribution is contained in a distributional uncertainty set [9]. In both cases, the goal is to find the best decision in view of the worst-

case realization of uncertainty. However, RO may propose overly conservative decisions than DRO since they do not exploit distributional information [24]. Both RO and DRO have wide application in many disciplines, such as energy, healthcare, transportation, logistics, and inventory [24]–[27]. Set-membership methods use a deterministic unknown-but-bounded description of noise and parametric uncertainties [28], [29]. Contrasted to such deterministic approaches, DRO is a stochastic approach in that uncertain parameters are assumed to follow some statistical distributions. For AMoD system balancing, Hao et al. consider the idle vehicle pre-allocation problem with uncertain demands and covariate information using DRO [6]; Miao et al. develop a data-driven DRO vehicle balancing method to accommodate uncertainties in the predicted demand distribution [30]. These methods only consider the demand uncertainty and cannot be directly applied to solve the challenge of integrally considering the demand and EV supply uncertainties. To the best of our knowledge, we are the first to consider both future mobility demand distribution uncertainties and EV supply distribution uncertainties.

III. Problem Formulation

Mobility demand uncertainty has been considered in AMoD vehicle allocation or balancing [1], [30]. EV supply uncertainty has also been addressed in EV charging scheduling [22], [23]. However, it is still challenging and has not been studied to simultaneously consider these two-sided uncertainties for EV AMoD systems, where the EV balancing for mobility and charging process is tightly integrated and should be optimized jointly.

In this section, we formulate the EV balancing problem as a distributionally robust optimization (DRO) problem considering both predicted passenger mobility demand and EV supply uncertainties. The DRO decision minimizes the worst-case expected cost over a set of uncertain supply-and-demand probability distributions and provides a performance guarantee of the decisions under model uncertainties. We consider both passenger supply-demand ratio fairness and EV charging supply-demand ratio fairness. While previous work considers only passenger mobility supply-demand ratio fairness in the objective function [30], we put it in the constraints and put EV charging supply-demand ratio fairness in the objective function (see Eq. (12) for detail). Such a problem formulation makes it possible to consider both charging and mobility fairness under two-sided uncertainties and decouple the mutual dependencies between EV supply and passenger demand under complex dynamics between EV charging and mobility patterns.

The balancing decisions are updated in a receding horizon control process [14], [31]. At each time step, the dispatching center first updates vehicle status and passenger demand information, then calculates the EV balancing decisions, and finally sends the decisions to the EVs to execute. Vacant EVs are allocated among different regions to pick up current and predicted future passengers. Low-battery EVs are dispatched to regions with charging stations to charge. The dispatching center focuses on the global-level balancing among regions. A local controller finishes one-to-one or one-to-group (carpool) EV-passenger or EV-charging station matching and detailed routing. We focus on the system-level robust EV balancing method design within a city. Local-level trip assignment and routing algorithms are out of the scope of this work and are investigated in the literature [32]–[34]. Our system-level EV balancing method can be applied in conjunction with these local-level one-to-one or one-to-group matching algorithms.

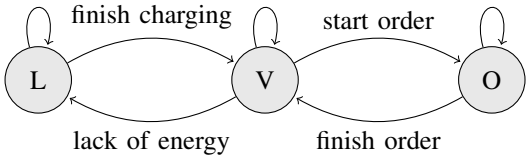


Fig. 2. Status transition process: the status "vacant" is a bridge between status "low-battery" and "occupied".

Nomenclature

Parameters of Problem (12)

K	number of time intervals in a day
N	number of regions
τ	model predicting time horizon
$r \in \mathbb{R}^{N\tau}$	concatenated demand vector with unknown distribution function F_r^*
$c \in \mathbb{R}^{N\tau}$	concatenated newly supply vector with unknown distribution function F_c^*
$W \in \mathbb{R}^{N\times N}$	cost matrix for vacant EVs, w_{ij} is the moving cost from region i to j for one vacant EV
$W^* \in \mathbb{R}^{N\times N}$	cost matrix for low-battery EVs, w_{ij}^* is the moving cost from region i to j for one low-battery EV
$P_v^k, P_o^k, P_l^k, Q_v^k, Q_o^k$	region transition matrices from time k to $(k+1)$
$V^1 \in \mathbb{N}^N$	the initial number of vacant EVs at each region provided by streaming data
$O^1 \in \mathbb{N}^N$	the initial number of occupied EVs at each region provided by streaming data
$m_1, m_2 \in \mathbb{R}_+$	upper bound of distance each EV can drive idly
$a \in \mathbb{R}_+$	power on the denominator of the objective
$\theta, \beta \in \mathbb{R}_+$	weight factors of the objective function
$l, h \in \mathbb{R}^\tau$	lower and upper bounds of mobility supply-demand ratio

Variables of Problem (12)

$X^k \in \mathbb{R}_+^{N\times N}$	x_{ij}^k is the number of vacant EVs dispatched from region i to region j at time k
$Y^k \in \mathbb{R}_+^{N\times N}$	y_{ij}^k is the number of low-battery EVs dispatched from region i to region j at time k
$V^k \in \mathbb{R}_+^N$	number of vacant EVs at each region before dispatching at the beginning of time k
$O^k \in \mathbb{R}_+^N$	number of occupied EVs at each region before dispatching at the beginning of time k
$L^k \in \mathbb{R}_+^N$	number of low-battery EVs at each region before dispatching at the beginning of time k
$S^k \in \mathbb{R}_+^N$	number of vacant (available) EVs at each region after dispatching at time k
$D, U \in \mathbb{R}_+^{N\tau}$	slack variables in constraint (8)

A. EV Status and Corresponding Actions

We define three statuses for one EV according to its battery level and working state: vacant, occupied, and low-battery. A vacant EV: one is working and has a battery level higher than a threshold **e**, but has no passengers. An occupied EV: one

is working and has a battery higher than a threshold **e** and has passengers. A low-battery EV: one is working and has no passengers in it and has a battery level lower than the threshold **e**. The controller dispatches vacant EVs according to current and predicted future passenger demands, and assigns low-battery EVs to regions where located charging stations. It has no actions for occupied EVs since these EVs are busy serving passengers. An EV's status can transit among these three categories. Fig. 2 shows the status transition process of the EVs according to the EV AMoD system dynamics [11], [34]–[37]. A low-battery/occupied EV can only transfer to a vacant EV or stay in the current status. The status "vacant" is a bridge between the status "low-battery" and "occupied".

B. DRO EV Balancing Problem for Mobility and Charging

Our goal is to find robust EV balancing decisions when considering the randomness and prediction errors of both EV supply and passenger mobility demand. Hence, we formulate a distributionally robust optimization (DRO) problem to minimize the expected EV balancing cost and provide fair charging and mobility service over an uncertainty set of demand and supply probability distributions. The nomenclature provides an overview of parameters and variables in the problem formulation and algorithm.

We divide one day into K time intervals and denote $k \in \{1, 2, ..., K\}$ as the temporal index. We separate a city into N regions and denote $i \in \{1, 2, ..., N\}$ as the spatial index. We denote r_i^k as the predicted total number of passengers demand, and c_i^k as the predicted total number of vacant EVs that finish charging and turn to supply in region i at time interval k. We then define $r^k = [r_1^k, r_2^k, ..., r_N^k]^\top \in \mathbb{R}^N$ and $c^k = [c_1^k, c_2^k, ..., c_N^k]^\top \in \mathbb{R}^N$ as vectors containing the demand and supply of each region in time interval k, $r = [r^1, r^2, ..., r^\tau] \in \mathbb{R}^{N\tau}$ and $c = [c^1, c^2, ..., c^\tau] \in \mathbb{R}^{N\tau}$ as the concatenation of demand and supply respectively. To consider model uncertainties, we assume they are random vectors instead of deterministic vectors. We use F_r^* and F_c^* to denote the unknown true probability distributions of $r, c \in \mathbb{R}^{N\tau}$ respectively, i.e., $r \sim F_r^*$ and $c \sim F_c^*$. We use non-negative matrices $X^k, Y^k \in \mathbb{R}^{N\times N}$ as the decision matrices, where x_{ij}^k, y_{ij}^k is the total number of vacant, low-battery EVs that will be dispatched from region i to region j at the beginning of time interval k, respectively. Minimizing the expected vacant and low-battery EVs balancing cost under random demand vector r and supply vector c is described as the following stochastic programming (SP) problem:

$$\begin{aligned} \min_{X^{1:\tau}, Y^{1:\tau}} \quad & \mathbb{E}_{r\sim F_r^*, c\sim F_c^*}\left[J(X^{1:\tau}, Y^{1:\tau}, r, c)\right] \\ \text{s.t.} \quad & X^{1:\tau}, Y^{1:\tau} \in \mathcal{D}, \end{aligned} \tag{1}$$

where $J(X^{1:\tau}, Y^{1:\tau}, r, c)$ is a cost function of allocating EVs given balancing decisions $X^{1:\tau} = \{X^1, X^2, \ldots, X^\tau\}$ and $Y^{1:\tau} = \{Y^1, Y^2, \ldots, Y^\tau\}$. $\mathcal{D}$ defines the convex constraints domain of the decision variables.

However, in the real world, we usually have limited knowledge about the true probability distributions F_r^* and F_c^*. With historical or streaming data, we can estimate sets of probability distributions $\mathcal{F}_r$ and $\mathcal{F}_c$ such that $F_r^* \in \mathcal{F}_r$, $F_c^* \in \mathcal{F}_c$ [12], [38], [39], instead of the exact F_r^* and F_c^*. Meanwhile, problem (1) is computationally expensive to solve in real-time [25]. Hence, we consider minimizing the worst-case expected cost, i.e., a minimax form of problem (1):

$$\begin{aligned} \min_{X^{1:\tau}, Y^{1:\tau}} \quad \max_{F_r\in\mathcal{F}_r, F_c\in\mathcal{F}_c} \quad & \mathbb{E}\left[J(X^{1:\tau}, Y^{1:\tau}, r, c)\right] \\ \text{s.t.} \quad & X^{1:\tau}, Y^{1:\tau} \in \mathcal{D}. \end{aligned} \tag{2}$$

Problem (2) is a distributionally robust optimization (DRO) problem [9], [25] that assumes $F_r^* \in \mathcal{F}_r$, $F_c^* \in \mathcal{F}_c$. In the following context, we define the complete forms of objective function and constraints. The formal definitions and construction algorithms of the distributional uncertainty sets $\mathcal{F}_r$ and $\mathcal{F}_c$ are introduced in Section IV.

Remark (Difficulties of considering multiple mutual effected parameter uncertainties)**.** *DRO methods have been designed in the literature [30] for vehicle balancing problems for gasoline vehicle MoD systems, where only the passenger mobility demand uncertainty has been considered. In contrast, our work considers the uncertainties of multiple parameters, i.e. demand and supply uncertainties. It is not straightforward to apply DRO methods that only consider a single parameter uncertainty to solve problems with multiple uncertain parameters, especially in a complicated and dynamic transportation system. The uncertain parameters of the system model are usually dynamically coupled [40], [41]. This mutual-dependent property results in two-fold difficulties. One is that formulating a computationally tractable DRO problem for vehicle balancing of an EV MoD system considering both the charging and passenger service processes gets more challenging. Considering the intrinsic connections of uncertain parameters, the DRO problem formulation should be carefully designed to satisfy necessary conditions of computational tractability, which is even challenging with a single uncertain parameter as shown in the literature [25]. To keep the problem formulation convex over decision variables and concave over uncertain parameters, we define the system performance requirements such as charging fairness in the objective function as a fractional function, and the mobility service fairness in the constraint functions as linear inequalities. We define the mobility dynamic process and status transition process of EVs as linear constraints. We explain all the detail in the problem formulation section. Another challenge is that numerically solving a DRO problem like* (12) *can be challenging given the definitions of the parameter uncertainty sets [25], [40], [41]. In this work, we decouple the dependencies of uncertain parameters in the problem formulation* (12)*, and then derive a computationally tractable and equivalent convex optimization problem for* (12) *in Section V.*

C. Balancing Cost

One optimization goal is to minimize the worst-case EV balancing cost. We define $W \in \mathbb{R}^{N\times N}$ as the cost matrix where w_{ij} is the cost of sending a vacant EV from region i to region j. The cost can be the approximated distance given a specific region partition method, approximated routing distance or travel time between two regions, etc. We select the approximated routing distance as the cost definition in our

data-driven experiments. We also define $W^* \in \mathbb{R}^{N\times N}, w^*_{ij}$ as the cost matrix of sending a low-battery EV. We set $w^*_{ij} = \infty$ if there are no charging stations in region j, since low-battery EV should not go to regions without charging stations.

Then the total balancing cost function J_D for τ intervals is defined as (3), where β is a positive weight coefficient.

$$J_D(X^{1:\tau}, Y^{1:\tau}) := \sum_{k=1}^{\tau}\sum_{i=1}^{N}\sum_{j=1}^{N}(x^k_{ij}w_{ij} + \beta y^k_{ij}w^*_{ij}). \quad (3)$$

D. Constraints Definition

1) Mobility Dynamics Constraints: We define the mobility dynamics constraints in (4) which describes the EV status transition. Let $V^k_i, O^k_i \in \mathbb{R}$ be the number of vacant and occupied EVs in region i at the beginning of time k before balancing, respectively. We define S^k_i as the total number of vacant EVs that are available to serve in region i after executing balancing decisions at time k. Then the following Equations (4) of V^k_i, O^k_i, S^k_i describe the EV AMoD dynamics [11], [34]–[37], [42] for $k \in \{1, \dots, \tau-1\}$:

$$\begin{aligned} S^k_i &= \sum_{j=1}^{N} x^k_{ji} - \sum_{j=1}^{N} x^k_{ij} + V^k_i, k \in \{1,\dots,\tau\};\\ V^{k+1}_i &= \sum_{j=1}^{N} P^k_{vji}S^k_j + \sum_{j=1}^{N} Q^k_{vji}O^k_j + c^k_i;\\ O^{k+1}_i &= \sum_{j=1}^{N} P^k_{oji}S^k_j + \sum_{j=1}^{N} Q^k_{oji}O^k_j, k \in \{1,\dots,\tau-1\}; \end{aligned} \quad (4)$$

where $P^k_v, P^k_o, Q^k_v, Q^k_o \in \mathbb{R}^{N\times N}$ are region transition matrices: $P^k_{vji}(P^k_{oji})$ is the probability that a vacant EV moves from region j at the beginning of time k will transverse to region i and being vacant (occupied) at the beginning of time $k+1$. Similarly, $Q^k_{vji}(Q^k_{oji})$ is the probability that an occupied EV moves from region j at time k will go to region i and being vacant (occupied) at the beginning of time $k+1$. The mobility constraints (4) show that (i) the number of vacant EVs after balancing is related to the number of vacant EVs before balancing and the net change number of vacant EVs according to the balancing decisions; (ii) there are two sources of occupied EVs: one is former occupied EVs that are still occupied, and the other is former vacant EVs that change into occupied; (iii) similarly, there are two sources of low-battery EVs: former low-battery EVs that stay in low-battery status, and former vacant EVs that change into low-battery status. The method of calculating region transition matrices is introduced in the literature [14]. When receding the time horizon, locations and status of all EVs are updated by real-time sensing data therefore V^1, O^1 are real-timely provided.

2) Charging Dynamics Constraints: We define L^k_i as the total number of low-battery EVs in region i before balancing at the beginning of time k, the charging dynamics constraints in Equation (5) states the quantitative relationship between low-battery EVs and available vacant EVs.

$$L^{k+1}_i = \sum_{j=1}^{N} y^k_{ji} - \sum_{j=1}^{N} y^k_{ij} + \sum_{j=1}^{N} P^k_{lji}S^k_j, \quad (5)$$

where L^1 is given by real-time data, $P^k_l \in \mathbb{R}^{N\times N}$ is the region transition matrix. Here P^k_{lji} is the probability that a vacant EV moves from region j at the beginning of time k will go to region i and being low-battery at the beginning of time $k+1$. The region transition matrices estimated from data satisfy that $\sum_{j=1}^{N} P^k_{lij} + P^k_{vij} + P^k_{oij} = 1$ and $\sum_{j=1}^{N} Q^k_{oij} + Q^k_{vij} = 1$. The method of calculating region transition matrices is introduced in the literature [14].

3) Moving Constraints: We also have moving constraints for decision variables X^k and Y^k defined in (6). When the balancing cost w_{ij} stands for distance, the idle driving distance EVs can move during a given time interval is limited, either due to speed limit or insufficient battery. Hence,

$$\begin{aligned} x^k_{ij} \geq 0 \text{ and } x^k_{ij} = 0 \text{ when } w_{ij} \geq m_1;\\ y^k_{ij} \geq 0 \text{ and } y^k_{ij} = 0 \text{ when } w^*_{ij} \geq m_2, \end{aligned} \quad (6)$$

where $m_1 > 0$ and $m_2 > 0$ is the upper bound moving distance for one vacant EV, and one low-battery EV, respectively. The values of m_1 and m_2 can be obtained by applying the comprehensive investigation results and methods in mobility and charging patterns of EVs from the literature [4], [43].

E. Mobility Supply-Demand Ratio

Mobility supply-demand ratio fairness is one service quality metric for AMoD systems [1], [15], [30]. In the literature, this goal is usually defined as minimizing the total absolute difference between local and global mobility supply-demand ratio for τ time intervals, i.e., $J_M(X^{1:\tau})$ defined in (7).

$$J_M(X^{1:\tau}) := \sum_{k=1}^{\tau}\sum_{i=1}^{N}\left|\frac{r^k_i}{S^k_i} - \frac{\sum_{j=1}^{N} r^k_j}{\sum_{j=1}^{N} S^k_j}\right|. \quad (7)$$

Note that the uncertain parameter c^k_i is related to S^k_i as defined in (4), the uncertain parameters r^k_i and c^k_i are included in the numerator and denominator of mobility demand-supply ratio, respectively. Hence, directly minimizing $J_M(X^{1:\tau})$ is computationally intractable. Instead of minimizing $J_M(X^{1:\tau})$ in (7), we consider the following mobility fairness constraints to make sure the mobility supply-demand ratio of each region is within a same range to provide fair service: $l^k \leq r^k_i/S^k_i \leq h^k, \quad k \in \{1,\dots,\tau\}$, where $l^k(h^k)$ is the lower (upper) bound of the mobility supply-demand ratio at time k. The value of l^k and h^k can be decided by historical data. We transfer those inequalities to the following Equations form with slack variables D^k_i, U^k_i:

$$r^k_i - l^k S^k_i - (D^k_i)^2 = 0,\ r^k_i - h^k S^k_i + (U^k_i)^2 = 0, \quad (8)$$

F. Charging Supply-Demand Ratio

When a fixed number and locations of charging stations are given, to avoid long waiting times or queues at some charging stations, we send low-battery EVs to regions with charging stations according to the dynamic availability of charging spots at different regions. Hence, we balance the charging supply-demand ratio for EV charging across the whole city [44].

We denote T_i^k as the net number of low-battery EVs in region i after the low-battery EV balancing decision Y^k,

$$T_i^k = \sum_{j=1}^{N} y_{ji}^k - \sum_{j=1}^{N} y_{ij}^k, \ \forall \ i \in \sigma. \tag{9}$$

where σ is the set of regions with charging stations. And we set $T_i^k = 0, \ \forall \ i \notin \sigma$, since low-battery EVs should not be dispatched to regions without charging stations. The charging supply-demand ratio in region i during time k is approximated as the ratio of the new charging spots supply and the EV charging demand rate c_i^k/T_i^k, where c_i^k is the number of EV that finished charging and become vacant vehicle supply in region i during time k. When one EV finishes charging, there will be one newly available charging spot; hence, we use c_i^k to approximate the amount of charging supply to serve low-battery EVs. To balance the charging supply-demand ratio over the city, we minimize the total absolute difference between the local and global charging supply-demand ratio for τ time intervals, i.e. J_1 defined in (10). Smaller $J_1(Y^{1:\tau})$ means the charging decisions $Y^{1:\tau}$ achieve higher charging fairness. Such fairness metrics have been widely used in literature [14], [30].

$$J_1(Y^{1:\tau}) = \sum_{k=1}^{\tau} \sum_{i \in \sigma} \left| \frac{c_i^k}{T_i^k} - \frac{\sum_{j \in \sigma} c_j^k}{\sum_{j \in \sigma} T_i^k} \right|. \tag{10}$$

However, function (10) is not concave over uncertainty parameter c_i^k. For computationally tractability, we define the objective of fair charging as to minimize the function J_E

$$J_E(Y^{1:\tau}) := \sum_{k=1}^{\tau} \sum_{i \in \sigma} \frac{c_i^k}{(T_i^k)^a}. \tag{11}$$

$J_E(Y^{1:\tau})$ has good properties that it is linear in c_k^i and convex over $Y^{1:\tau}$ when the power parameter $a > 0$. What's more, the function (11) approximates function (10) when a is designed to be small enough, according to Lemma 1 in the literature [30].

G. DRO EV Balancing Problem Formulation

With constraints (4), (5), (6), (8), we finally define the distributionally robust EV balancing for mobility and charging process under uncertain probability distributions of the demand and supply in (12), where the objective function J is a weighted sum of the balancing cost function J_D defined in (3) and charging unfairness function J_E defined in (11).

$$\begin{aligned} \min_{\substack{X^{1:\tau}, Y^{1:\tau}, S^{1:\tau}, D^{1:\tau}, \\ U^{1:\tau}, V^{2:\tau}, O^{2:\tau}, L^{2:\tau}}} \ \max_{\{F_r \in \mathcal{F}_r, F_c \in \mathcal{F}_c\}} & \ \mathbb{E}\left[J\right] \\ \text{s.t. (4); (5); (6); (8),} & \end{aligned} \tag{12}$$

where J is the final objective defined in (13),

$$J = J_D(X^{1:\tau}, Y^{1:\tau}) + \theta J_E(Y^{1:\tau}, c) \tag{13}$$

θ is a positive weighted parameter for a trade-off between J_D and J_E. And $T_i^k = \mathcal{L}_T(Y^k)$ is a linear function of decision variables Y^k, $S_i^k = \mathcal{L}_S(X^{1:k}, c_i^{1:k-1})$ is a linear function of decision variables $X^{1:k} := \{X^1, ..., X^k\}$ and supply uncertain parameter $c_i^{1:k-1} := \{c_i^1, ..., c_i^{k-1}\}$. We prove the following Lemma 1, there must exist a set of lower and upper bounds of the mobility supply-demand ratio in constraint (8) to guarantee that there is a feasible solution for the proposed DRO problem (12).

Lemma 1. *We can find at least one lower and one upper bound of the mobility supply-demand ratio such that the DRO problem* (12) *has at least one feasible solution. We name these bounds global bounds. And there exists a pair of lower and upper bounds l^k and h^k that are no looser than global bounds as well as guarantee the DRO problem* (12) *has at least one feasible solution.*

Proof. See Appendix VIII-A. □

Remark (Novelty of problem formulation)**.** *The DRO EV balancing problem formulation* (12) *decouples the mutual dependencies between EV supply and passenger demand. By putting the mobility supply-demand ratio fairness requirement in the constraints and the EV charging supply-demand ratio fairness in the objective, we make it possible to consider both charging and mobility fairness under two-sided uncertainties. We then derive a computationally tractable convex optimization form of problem* (12) *in Section V.*

H. Generalization for Heterogeneous EV Fleet

We generalize the DRO problem (12) with homogeneous EVs (i.e. all vacant EVs have the same capacity or number of available seats) to a heterogeneous EV balancing problem formulation that considers EVs with different capacities.

We denote the capacity of one EV as C_e, where e (type) is the index of different types as $e = 1, \cdots, E$. E is the total number of capacity types. For instance, C_e is the number of seats for type e EVs. We use $V_{e,i}^k$ to denote the number of vacant EVs with capacity C_e in region i at time k before balancing. So the total number vacant EVs in region i at time k before balancing is $V_i^k := \sum_{e=1}^{E} V_{e,i}^k$. Similarly, we use $S_{e,i}^k, O_{e,i}^k, L_{e,i}^k$ to denote numbers of available, occupied, low-battery EVs with C_e capacity, respectively. Then $x_{e,ij}^k$ is the number of type-e vacant EVs that will be dispatched from region i to region j in time interval k. And $y_{e,ij}^k$ is the decision for type-e low-battery EVs. Then the balancing cost constraints (6) become:

$$\begin{aligned} x_{e,ij}^k \geq 0 \text{ and } x_{e,ij}^k = 0 \text{ when } w_{ij} \geq m_{e,1}; \\ y_{e,ij}^k \geq 0 \text{ and } y_{e,ij}^k = 0 \text{ when } w_{ij}^* \geq m_{e,2}, \end{aligned} \tag{14}$$

because different types of EVs may have different moving constraints. For example, EVs with a smaller number of seats may use less energy or time to finish the same-length trip compared to EVs with a larger number of seats. Because they usually benefit from lighter cargo and metal frames, as well as fewer traffic constraints (due to height limit, size limit, lane limit, etc.). We can determine these constraints parameters by applying the comprehensive investigation results and methods in mobility and charging patterns of EVs from the

literature [4], [43]. And the mobility dynamic constraints (4) become: for $e = 1, \cdots, E$,

$$\begin{aligned} S_{e,i}^k &= \sum_{j=1}^N x_{e,ji}^k - \sum_{j=1}^N x_{e,ij}^k + V_{e,i}^k, \quad k = 1, \cdots, \tau; \\ V_{e,i}^{k+1} &= \sum_{j=1}^N P_{e,vji}^k S_{e,j}^k + \sum_{j=1}^N Q_{e,vji}^k O_{e,j}^k + c_{e,i}^k, \\ O_{e,i}^{k+1} &= \sum_{j=1}^N P_{e,oji}^k S_{e,j}^k + \sum_{j=1}^N Q_{e,oji}^k O_{e,j}^k, k = 1, \cdots, \tau - 1, \end{aligned} \tag{15}$$

where $S_{e,i}^k$ is the total number of type-e vacant EVs in region i at time k, $c_{e,i}^k$ is the number of type-e low-battery EVs in region i that finish charging in time k. The charging dynamic constraint (5) turns to:

$$L_{e,i}^{k+1} = \sum_{j=1}^N y_{e,ji}^k - \sum_{j=1}^N y_{e,ij}^k + \sum_{j=1}^N P_{e,lji}^k S_{e,j}^k. \tag{16}$$

The region transition matrices $P_{e,vij}^k, P_{e,oij}^k, P_{e,lij}^k$ denote the probability that an available type-e EV moves from region i at the beginning of time k will transverse to region j and being vacant, occupied, low-battery, respectively at the beginning of time $k+1$. Similarly, $Q_{e,vij}^k, Q_{e,oij}^k$ denote the probability that an occupied type-e EV becomes vacant, occupied, respectively in the transition process. The objective J_E has a new form that:

$$J_E' := \sum_{k=1}^{\tau} \sum_{i \in \sigma} \frac{\sum_{e=1}^E c_{e,i}^k}{(\sum_{e=1}^E T_{e,i}^k)^a}. \tag{17}$$

The objective J_D also has a new form that:

$$J_D' := \sum_{e=1}^E \sum_{k=1}^{\tau} \sum_{i=1}^N \sum_{j=1}^N (x_{e,ij}^k w_{ij} + \beta y_{e,ij}^k w_{ij}^*). \tag{18}$$

The objective function J_E' defined in Equation (17) is linear (concave) over c^k, convex over $Y_{1:E}^{1:\tau}$. The objective function J_D' defined in Equation (18) is convex over $X_e^{1:\tau}$ and $Y_e^{1:\tau}$, $e = 1, \cdots, E$, since linear operation preserves convexity [47]. The generalized balancing cost constraints (14), mobility dynamic constraints (15) and charging dynamic constraints (16) are linear of decision variables. So we have Lemma 2.

Lemma 2. *The generalized heterogeneous EV balancing problem* (19)

$$\begin{aligned} &\underset{\substack{X^{1:\tau}, Y^{1:\tau}, S^{1:\tau}, D^{1:\tau}, \\ U^{1:\tau}, V^{2:\tau}, O^{2:\tau}, L^{2:\tau}}}{\textit{min.}} \quad \underset{\{F_r \in \mathcal{F}_r, F_c \in \mathcal{F}_c\}}{\textit{max.}} \mathbb{E}\left[J_D' + \theta J_E'\right] \\ &\textit{s.t.} \quad (8); (14); (15); (16). \end{aligned} \tag{19}$$

is still convex with respect to the decision variables, and concave on uncertain parameters, where J_E' and J_D' are defined in (17) *and* (18) *respectively.*

Proof. See Appendix VIII-B. □

IV. Distributional Uncertainty Set Form and Construction Algorithm

We design an efficient algorithm to construct the demand and supply uncertainty sets $\mathcal{F}_r, \mathcal{F}_c$ of probability distributions defined in problem (12), with historical data that contains information related to the true demand and supply distributions. The works in [25] and [45] use empirical estimates to construct the uncertainty set according to confidence regions of hypothesis testing in portfolio management problems. The method in [30] leverages the structure property of the covariance matrix to develop efficient uncertainty sets construction algorithms. However, the literature methods do not provide stable uncertainty set guarantees. Hence, we design a bootstrap-based algorithm that produces stable uncertainty set given different prediction models, and prove that the produced uncertainty set parameters are guaranteed to be contained in desired confidence regions with a given probability.

A. General Distributional Uncertainty Set

Without loss of generality, we denote z as a random vector variable and denote $\{z_1, ..., z_n\}$ as a set of sample vectors; $f(o)$ as a prediction model of random variable z and $\hat{z}$ as the prediction of z by the model $f(o)$, i.e., $\hat{z} = f(o)$, where o is the input data. We use δ to denote the prediction error which is the difference between the true value and the estimated value, i.e., $z = \hat{z} + \delta$. Therefore, δ is also a random vector variable. We define the following distributional uncertainty set construction problem for random variable z. We then calculate the uncertainty sets for both random demand r and supply c according to formulation (20).

Problem 1. *Given a sample set of the random vector variable z, a prediction method $f(o)$, find values of $\hat{z}, \hat{\Sigma}, \hat{\omega}$ and $\hat{\gamma}$, such that with probability at least $1 - \alpha$ the true distribution of z is contained in the distributional uncertainty set* (20).

$$\begin{aligned} \mathcal{F}_z(\hat{z}, \hat{\Sigma}, \hat{\omega}, \hat{\gamma}) = \{z = \hat{z} + \delta : \\ \mathbb{E}[\delta]^T \hat{\Sigma}^{-1} \mathbb{E}[\delta] \leqslant \hat{\omega}, \mathbb{E}(\delta \delta^T) \preceq \hat{\gamma} \hat{\Sigma}\}, \end{aligned} \tag{20}$$

where $\hat{\Sigma}$ is the estimate covariance of δ, $\hat{\omega}$ and $\hat{\gamma}$ are two estimated constraints parameters.

Definition (20) means that the distributional uncertainty set $\mathcal{F}_z$ relies on the estimated covariance matrix of prediction error δ. The mean of δ is supposed to lie in an ellipsoid. The estimate covariance $\hat{\Sigma}$ lies in a positive semi-definite cone defined with a matrix inequality.

B. Uncertainty Set Construction Algorithm

We develop Algorithm 1 to solve Problem 1. Algorithm 1 computes the constraint parameters based on the bootstrap sample method [46]. Given a pre-determined prediction model, historical data and observation data, in the outer loop, we sample with replacement from the historical sample set to train prediction model, then compute the set of prediction residual $\tilde{\Delta}$, i.e., realizations of prediction error, compute the sample mean and sample covariance. Then we compute constraint parameters by solving (21) and (22).

$$\omega_j^i = [\bar{\delta}_j^i]^T (\hat{\Sigma}^i)^{-1} [\bar{\delta}_j^i] \tag{21}$$

$$\min_{\gamma_j^i} \gamma_j^i \quad \text{s.t} \quad \bar{\Sigma}_j^i \preceq \gamma_j^i \hat{\Sigma}^i, \tag{22}$$

where $\bar{\delta}_j^i$ is a sample mean; $\bar{\Sigma}_j^i$ is a sample covariance; $\hat{\Sigma}^i = \sum_{j=1}^N \bar{\Sigma}_j^i$ is the estimated covariance in the i-th outer

Algorithm 1: Uncertainty Set Parameters Estimation

Data: a set of samples Z, bootstrap time A, bootstrap sample set size M, inner-bootstrap time B and inner-bootstrap sample set size N, a significance level α, a confidence parameter η, a prediction model f and a observation set O.

Result: Uncertainty Set Parameters and Confidence Regions

while *i in 1 to A* **do**
  while *j in 1 to B* **do**
    Re-sample $Z_j^i = \{\tilde{z}_{j1}^i, ..., \tilde{z}_{jN}^i\}$ from Z with replacement, estimate parameters of a prediction model $f_j^i(O)$, calculate the estimation residual set $\tilde{\Delta}_j^i = \{\tilde{\delta}_{j1}^i, ..., \tilde{\delta}_{jN}^i\}$, where $\tilde{\delta}_{jk}^i = \tilde{z}_{jk}^i - \hat{z}_{jk}^i$. Compute its sample mean $\bar{\delta}_j^i$ and sample covariance $\bar{\Sigma}_j^i$.
  end
  Compute estimated covariance $\hat{\Sigma}^i = \frac{1}{B}\sum_{j=1}^{B} \bar{\Sigma}_j^i$.
  while *j in 1 to B* **do**
    Compute ω_j^i and γ_j^i according to (21) and (22) get $\Omega^i = \{\omega_1^i, ..., \omega_B^i\}$ and $\Gamma^i = \{\gamma_1^i, ..., \gamma_B^i\}$
  end
  Compute ω_α^i from Γ^i and γ_α^i from Γ^i
  while *k in 1 to C* **do**
    Sampling with replacement from set Ω^i and Γ^i to get bootstrap sample sets Ω^{ik} and Γ^{ik};
    Compute ω_α^{ik} from Ω^{ik} and γ_α^{ik} from Γ^{ik}
  end
end
Compute sample standard deviations: s_ω^i, s_γ^i, s_ω, s_γ, estimated parameters: $\hat{\omega}_\alpha$, $\hat{\gamma}_\alpha$, $\hat{\Sigma}$, quantiles: $q_{\eta/2}^\omega$ and $q_{1-\eta/2}^\omega$, $q_{\eta/2}^\gamma$ and $q_{1-\eta/2}^\gamma$

loop; $i \in \{1, 2, ..., A\}, j \in \{1, 2, ..., B\}$ denote the outer loop and inner loop numbers, respectively. Now we compute the α percentiles for the sets $\Omega^i = \{\omega_j^i\}_{j\in\{1,2,...,B\}}$, $\Gamma^i = \{\gamma_j^i\}_{j\in\{1,2,...,B\}}$, which we denote as $\omega_\alpha^i, \gamma_\alpha^i$ respectively. We then sample with replacement from Ω^i and Γ^i respectively to get constraint parameter sample sets Ω^{ik} and Γ^{ik} for C times, where $k \in \{1, ..., C\}$ denotes loop numbers. We also compute the α percentiles ω_α^{ik} for Ω^{ik}, γ_α^{ik} for Γ^{ik}.

Finally we get estimated constraint parameters $\hat{\omega}_\alpha$, $\hat{\gamma}_\alpha$ and estimated covariance $\hat{\Sigma}$ by (23).

$$\hat{\omega}_\alpha = \sum_{i=1}^{A} \omega_\alpha^i / A, \hat{\gamma}_\alpha = \sum_{i=1}^{A} \gamma_\alpha^i / A, \hat{\Sigma} = \sum_{i=1}^{A} \hat{\Sigma}^i / A. \tag{23}$$

Then we compute the $\eta/2$ and $1-\eta/2$ quantiles: $q_{\eta/2}^\omega$ and $q_{1-\eta/2}^\omega$ for $\{\frac{\omega_\alpha^i - \hat{\omega}_\alpha}{s_\omega^i}\}_{i\in\{1,...,A\}}$, $q_{\eta/2}^\gamma$ and $q_{1-\eta/2}^\gamma$ for $\{\frac{\gamma_\alpha^i - \hat{\gamma}_\alpha}{s_\gamma^i}\}_{i\in\{1,...,A\}}$, where s_ω^i and s_γ^i are standard deviations on the set of $\{\omega_\alpha^{ik}\}_{k\in\{1,...,C\}}$ and $\{\gamma_\alpha^{ik}\}_{k\in\{1,...,C\}}$, respectively. We define the confidence regions of ω_α and γ_α: $[\omega_l, \omega_u]$ and $[\gamma_l, \gamma_u]$, with the lower and upper bounds in (24),

$$\begin{aligned} \omega_l &= \hat{\omega}_\alpha - s_\omega q_{1-\eta/2}^\omega, \quad \omega_u = \hat{\omega}_\alpha - s_\omega q_{\eta/2}^\omega, \\ \gamma_l &= \hat{\gamma}_\alpha - s_\gamma q_{1-\eta/2}^\gamma, \quad \gamma_u = \hat{\gamma}_\alpha - s_\gamma q_{\eta/2}^\gamma, \end{aligned} \tag{24}$$

where s_ω and s_γ are respectively standard deviations on the set of $\{\omega_\alpha^i\}_{i\in\{1,...,A\}}$ and $\{\gamma_\alpha^i\}_{i\in\{1,...,A\}}$. These two confidence regions are guaranteed to contain the true constraint parameters ω_α and γ_α respectively, at the $1-\eta$ confidence level.

Lemma 3 provides the theoretical support to the computing procedure of confidence regions. Thus, the produced uncertainty set parameters are guaranteed to be contained in some confidence regions with a given probability.

Lemma 3. *Given a pre-selected probability $1-\eta$, parameters calculated by Algorithm 1 satisfy that ω_α is contained between $\hat{\omega}_\alpha - s_\omega q_{\eta/2}^\omega$ and $\hat{\omega}_\alpha - s_\omega q_{1-\eta/2}^\omega$, γ_α is contained between $\hat{\gamma}_\alpha - s_\gamma q_{\eta/2}^\gamma$ and $\hat{\gamma}_\alpha - s_\gamma q_{1-\eta/2}^\gamma$.*

Proof. See Appendix VIII-C □

V. Computationally Tractable Form

In this section, we derive the main theoretical result of this work: Theorem 1, a computationally tractable and equivalent convex optimization form for the distributionally robust optimization problem (12) via strong duality. The objective function (13) is convex over the decision variables and linear (concave) over the random parameter, with decision variables on the denominators. Though constraints are affine or convex over decision variables, most of them contain random parameters. This form is not a Linear Programming (LP) [47] or Semidefinite Programming (SDP) [48]. The main process of deriving an equivalent convex optimization problem is to analyze (13) and the constraints part. Based on Theorem 1, the optimal solution to (12) can be calculated in real-time considering both passenger demand and EV supply uncertainties.

Theorem 1. *The distributionally robust optimization problem* (12) *with two distributional sets* (20) *is equivalent to the following convex optimization problem*

$$\begin{aligned} \min_{\substack{X^{1:\tau}, Y^{1:\tau}, D^{1:\tau}, U^{1:\tau}; \\ S^{1:\tau}, V^{2:\tau}, O^{2:\tau}, L^{2:\tau}; \\ \lambda, Q_r, q_r, v_r, t_r, Q_c, q_c, v_c, t_c}} & \quad H_o + v_r + t_r + v_c + t_c \\ \text{s.t.} \quad & \begin{bmatrix} v_r & \frac{1}{2}(q_r + \lambda_U + \lambda_D)^T \\ \frac{1}{2}(q_r + \lambda_U + \lambda_D) & Q_r \end{bmatrix} \succeq 0, \\ & \begin{bmatrix} v_c & \frac{1}{2}(q_c + \lambda_V - Z)^T \\ \frac{1}{2}(q_c + \lambda_V - Z) & Q_c \end{bmatrix} \succeq 0, \\ & t_r \geqslant (\hat{\gamma}_r \hat{\Sigma}_r + \hat{r}\hat{r}^T) \cdot Q_r + \hat{r}^T q_r \\ & \quad + \sqrt{\hat{\omega}_r} \|\hat{\Sigma}_r^{1/2}(q_r + 2Q_r\hat{r})\|_2, \\ & t_c \geqslant (\hat{\gamma}_c \hat{\Sigma}_c + \hat{c}\hat{c}^T) \cdot Q_c + \hat{c}^T q_c \\ & \quad + \sqrt{\hat{\omega}_c} \|\hat{\Sigma}_c^{1/2}(q_c + 2Q_c\hat{c})\|_2, \\ & Q_r, Q_c \succeq 0, \qquad \lambda, v_r, v_c \geq 0, \\ & z_i^k \geqslant \frac{1}{(T_i^k)^a} \\ & x_{ij}^k \geqslant 0 \text{ and } x_{ij}^k = 0 \text{ when } w_{ij} \geq m_1; \\ & y_{ij}^k \geqslant 0 \text{ and } y_{ij}^k = 0 \text{ when } w_{ij}^* \geq m_2, \end{aligned} \tag{25}$$

where $H_o = J_D - (\lambda_S^T f_S + \lambda_O^T f_O + \lambda_L^T f_L + \lambda_s^T S^{1:\tau} + \lambda_l^T L^{2:\tau}) - \lambda_D^T diag(lS^T - DD^T) - \lambda_U^T diag(hS^T + UU^T) - \sum_{k=1}^{\tau-1}(-V_i^{k+1} + \sum_{j=1}^{N} P_{vji}^k S_j^k + \sum_{j=1}^{N} Q_{vji}^k O_j^k)\lambda_{V_i^k}$, J_D *is defined as* (3).

Proof. See Appendix VIII-D. □

This convex optimization problem (25) has a objective which is a linear function of decision variables v_r, t_r, v_c, t_c, $X^{1:\tau}$, $Y^{1:\tau}$, $S^{1:\tau}$, $V^{2:\tau}$, $O^{2:\tau}$, $L^{2:\tau}$ and λ; as well as a quadratic function of decision variables $D^{1:\tau}$ and $U^{1:\tau}$. It includes linear matrix inequalities constraints, linear norm inequalities constraints and other convex constraints.

Remark (Computational Complexity)**.** *In the DRO EV balancing problem, every uncertain parameter is in a convex and compact support set. Each support set is equipped with an oracle which has the following properties. First, given any value of the uncertain parameter, within polynomial time (with respect to the dimension of the uncertain parameter), the oracle can (i) either confirm the value is within the support set; (ii) or provide a hyperplane that separates the value from the support set [49]. Second, given all parameters and variables, the value of* J_D *and* J_E *can be evaluated in polynomial time. Notice that the problem* (25) *has a linear objective function and convex constraint functions over decision variables. Hence, Proposition 1 in [25] can be applied and the problem* (25) *can be solved to any precision* ϵ *within polynomial time of* $\log(1/\epsilon)$ *and the size of the problem. This equivalent computationally tractable convex optimization problem can be solved by existing convex optimization problem solvers such as CVXPY [50]. For instance, CVXOPT converts a problem into its equivalent standard form known as conic form, and provide polynomial interior-point algorithms [51] to solve it.*

VI. Experiment

In this section, we evaluate the performance of the designed distributionally robust optimization-based EV balancing method with electric taxi (e-taxi) data from the Chinese city Shenzhen (one of the largest cities in China, which operates over 10,000 E-taxis). Six-week real-world data is utilized in the experiment, which includes 60GB EV GPS data, and 5.5 GB transaction data from over 10,000 EVs. We split the whole dataset 67%-33% between the training set and the testing set. This 2:1 ratio is commonly used in train-test split for evaluating machine learning algorithms. Even though two weeks seem short for testing compared to the data used in other machine learning methods, the data size is large (20GB GPS and 1.9 GB transaction data involving over 10,000 EVs). To our knowledge, few existing EV works utilized such large-scale data from large-scale EV deployments for algorithm evaluation. In Fig. 1, there is a high-level illustrative flow graphic to describe the proposed solution framework to help readers understand the experimental procedures.

A. Data Description

In total, there are four different datasets used in this paper, including E-taxi GPS data (vehicle ID, locations, time and speed, etc), Transaction data (vehicle ID, pick-up and drop-off time, pick-up and drop-off location, travel distance, etc), charging station data (locations, name, the number of charging ports, etc), and city partition data (geographic boundaries of 491 small separate regions composing Shenzhen). Our datasets are obtained by collaborating with the Shenzhen Transportation Committee for its smart city initiative. An example of these four datasets is shown in the Tab. I.

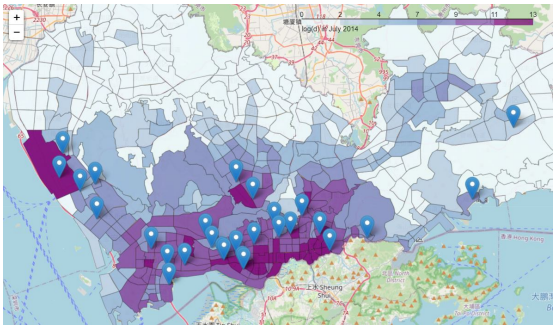


Fig. 3. Heat map of demand in Shenzhen City: lighter means less demand.

GPS Data include fields of the status of taxis, e.g., the vehicle ID, the GPS location longitude and latitude information, time-stamp, direction, current speed, etc. **Transaction Data** describe each transaction information, e.g., vehicle ID, pick-up/drop-off time, pick-up/drop-off location (longitude & latitude), and travel distance, etc. **Charging Station Data** include the locations (i.e., GPS) of each charging station, station name, and the number of charging ports in each station, etc. **Urban Partition Data** describe the urban partition for population census of the Shenzhen city. There are 491 regions in total, and each region has a region ID and longitudes & latitudes of its boundary.

Fig. 3 is a heat map of demand in Shenzhen city. It shows the unbalanced distribution of mobility demand. There is higher passenger demand in downtown and airport areas than in suburban areas. This realistic situation verifies the necessity to provide fair service in EV AMoD systems. These blue markers in Fig. 3 denote charging station locations. The unbalanced charging stations distribution also enhances our motivation to consider charging fairness in EV AMoD balancing problems.

B. Data Processing

For the host machine and software used to clean and manipulate the original datasets, we utilize a 34 TB Hadoop Distributed File System (HDFS) on a cluster consisting of 11 nodes, each of which is equipped with 32 cores and 32 GB RAM. To make this data fit within the context of our problem formulation, we further clean our datasets according to several principles including continuity of time, homogeneity in data distributions, uniform time granularity, etc. EV mobility and charging patterns are also considered in the filtering process and a comprehensive investigation of the process has been introduced in the literature [52].

To merge multi-source datasets, we first select the electric taxis with complete GPS and the transaction records in the dataset. Then we insert the departure and arrival of each transaction as new GPS timestamps and have area information labeled, building an aligned timeline for each taxi. We

TABLE I
AN EXAMPLE OF THE FOUR DATASETS

GPS Data	plate ID	longitude	latitude	time	speed (km/h)
	BDXXXX	114.0121	22.526104	2015-08-16 08:30:43	35
Transaction Data	plate ID	pickup time	dropoff time	pickup location	travel distance (m)
	BDXXXX	2015-09-03 13:47:58	2015-09-03 13:57:23	(113.9867, 22.5433)	6954
Charging Station Data	station ID	station name	longitude	latitude	number of charging ports
	30	NB0005	113.9878608	22.55955418	40
Urban Partition Data	Region ID	Longitude1	Latitude1	Longitude2	Latitude2
	1	114.31559657	22.78559093	114.311230763	22.78220351

further utilize a widely adopted spatiotemporal constraint-based method in the literature [12], [38] to infer the charging activities taken by taxis based on their individual timeline. We also define 7 mutually exclusive labeling codes for marking timestamps' riding or charging status, based on which we can obtain the statistics of EV activities in both spatial and temporal dimensions.

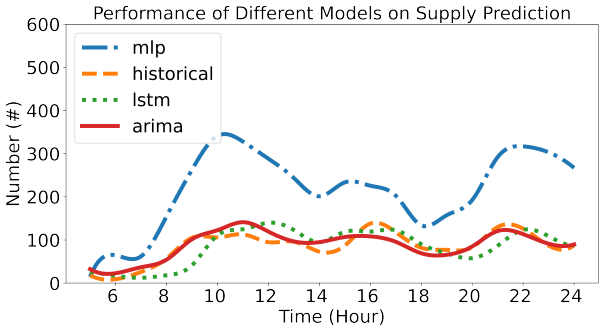


Fig. 4. The performance of ARIMA model in predicting supply and demand is better than MLP and LSTM models. So we decide ARIMA model as the prediction model f when using Algorithm 1 to construct uncertainty sets.

C. Prediction Model

We decide which prediction model to be used by comparing the performance of these three models: Long Short-Term Memory model (LSTM) [53], AutoRegressive Integrated Moving Average model (ARIMA) [54] and Multi-Layer Perceptron model (MLP) [55]. In Fig. 4, the three models all show the general trend of supply in one day. But ARIMA model performs better than other two models in term of prediction accuracy. In Fig. 5, the prediction errors for supply at each hour are reported. MLP model has the largest prediction errors since the blue curve is almost always higher than all the other curves. The curve of LSTM model is close to the ground truth (the zero horizon line) but still has larger errors than the ARIMA model. To quantitatively compare the performance of these models, we further provide the mean square error (MSE) in Tab. III. MSE is the average squared difference between

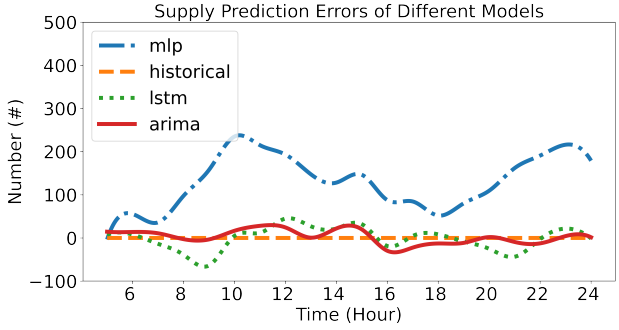


Fig. 5. The ARIMA model has the lowest supply prediction errors.

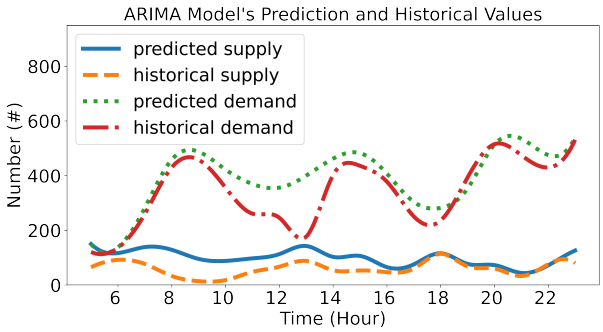


Fig. 6. ARIMA model demonstrates time trends very well.

the predicted values and the historical values. It measures the prediction quality of a model and a lower MSE indicates a better accuracy. We can see that ARIMA model achieves the smallest MSE. Hence, we decide ARIMA model as the prediction model f when using Algorithm 1 to construct uncertainty sets. In Fig. 6, we compare ARIMA model's predicted values and historical values of supply and demand on one day in region 2. ARIMA model demonstrates supply and demand's time trends very well and the predicted values of demand in peak hours: 8am-10am, 2pm-4pm, 8pm-10pm, are close to historical values.

D. Distributional Uncertainty Set

Tab. II shows how the constraint parameters $\hat{\omega}_c$, $\hat{\gamma}_c$ change for different parameter B of Alg. 1. The values of $\hat{\omega}_c$ and $\hat{\gamma}_c$ decrease when B increases and the speed of decreasing turns slower as B becoming larger. We also notice that as B increases, the corresponding estimated constraint parameters $\hat{\omega}_c$ and $\hat{\gamma}_c$ have a trend to converge to a certain constant that meets bootstrapping algorithm's intuition. Here one time interval is 1 hour, the time horizon τ is 2, the significant $\alpha = 0.25$ and the true probability distributions of demand and supply variables r, c are separately contained in the constructed distributional uncertainty sets with probability 75%.

TABLE II
THRESHOLDS $\hat{\omega}_c$ AND $\hat{\gamma}_c$ FOR DIFFERENT PARAMETERS B

B	10	20	50	100	500	1000
$\hat{\omega}_c$	1.504	0.964	0.576	0.399	0.296	0.176
$\hat{\gamma}_c$	3.715	2.832	2.006	1.768	1.374	1.317

We validate the robustness of our proposed uncertainty set construction Alg. 1 in Fig. 7-1 and Fig. 7-2. We run Alg. 1 with different bootstrap time A and the baseline algorithm [30] for 20 times, using the same prediction model and parameters. We use $A = 0$ to denote the baseline algorithm. Then we compute the variances of the outputted uncertainty set parameters. Variance is an indicator of the robustness of random algorithms [56]. It is known that in several general cases, a random

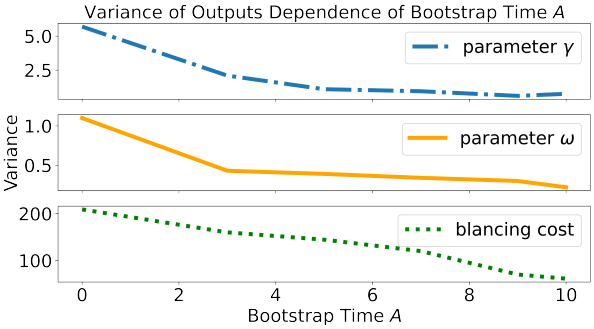


Fig. 7. The variances of constraint parameters/balancing cost are decreased as the bootstrap time A increasing.

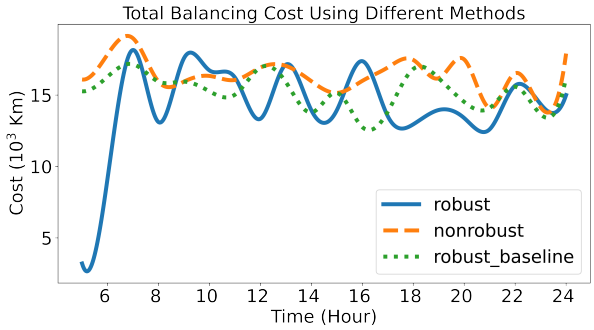


Fig. 8. By using our DRO method, the average total balancing cost is reduced by 14.49% compared to non-robust method.

algorithm which produces a set of outputs with a smaller variance is more robust. Compared to the baseline algorithm, our Alg. 1 produces outputted constraint parameters with smaller variances. In particular, our construction algorithm improve the robustness to randomness by at least 65%, compared with the baseline algorithm. And the values of variance decrease as A getting larger.

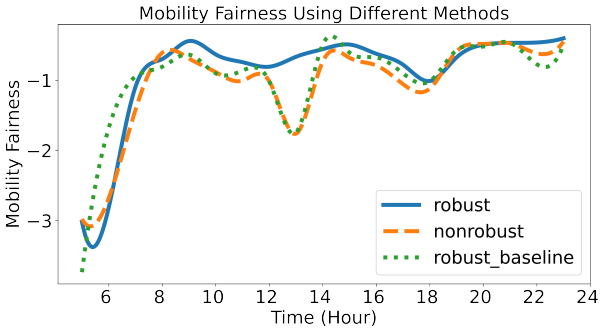


Fig. 9. By using our DRO method, the average fairness of mobility supply-demand ratio is improved by 15.78% compared to non-robsut method.

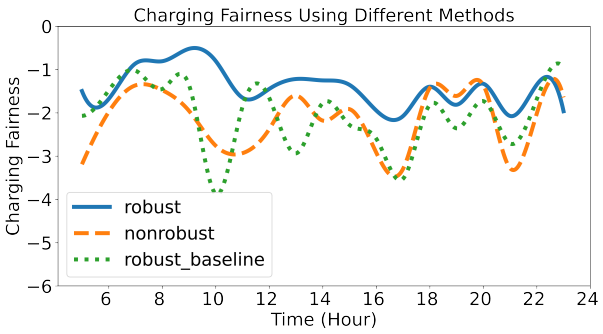


Fig. 10. By using our DRO method, the average fairness of charging supply-demand ratio is improved by 34.51% compared to non-robsut method.

TABLE III
MEAN SQUARE ERROR OF DIFFERENT MODELS

	MLP	ARIMA	LSTM
MSE	20148.21	254.65	647.45

E. Performance Comparison

We compare our distributionally robust optimization (DRO) method with the baseline robust method [6] and the non-robust method [57] by using the same real-time sensing data. The non-robust method is described in Appendix, please refer to Equation (41). The EV balancing decision is made every hour, and the running time of our algorithm is within 5 seconds (3.76 s $\pm$1.18), which can satisfy the real-time requirement of making balancing decisions. Fig. 8 shows the total balancing cost M_b^k from 5am to 11pm using our method (12) as well as two baseline methods. The metric M_b^k is formally given in (26), which is a weighted sum of the vacant and low-battery EVs' moving distance after executing balancing decisions at time k. A smaller M_b^k is better, since a smaller total balancing cost indicates higher system efficiency and lower distance that the EVs run without serving passengers.

$$M_b^k = J_D(X^k, Y^k) = \sum_{i=1}^{N}\sum_{j=1}^{N}(x_{ij}^k w_{ij} + \beta y_{ij}^k w_{ij}^*). \quad (26)$$

We can see that most of the time, the total cost of our method is lower than that of the two baseline methods. In particular, the average total balancing cost is reduced by 14.49% compared with the non-robust method and reduced by 7.37% compared with the baseline robust method. In Fig. 7-3, we also show the variance of the optimal daily balancing cost of the DRO (12) solutions when using Alg. 1 with different bootstrap parameters A and the baseline uncertainty set construction algorithm [30] (A = 0). For an uncertainty set produced in different settings, we keep other parameters the same and run our DRO method for 20 times. Then we compute the variance of daily balancing cost $\sum_{k=1}^{K} M_b^k$. The variance of the daily balancing cost decreases as A gets larger. And when using our Alg. 1, we obtain a lower variance of daily balancing cost, compared with the baseline algorithm. In particular, the variance is reduced by at least 31%, compared with the baseline algorithm. Our DRO method achieves more stable balancing cost, by using together with our proposed uncertainty set construction algorithm.

$$M_c^k = -J_1(Y^k) = -\sum_{i\in\sigma}\left|\frac{c_i^k}{T_i^k} - \frac{\sum_{j\in\sigma} c_j^k}{\sum_{j\in\sigma} T_i^k}\right|; \quad (27)$$

$$M_m^k = -J_M(X^k) = -\sum_{i=1}^{N}\left|\frac{r_i^k}{S_i^k} - \frac{\sum_{j=1}^{N} r_j^k}{\sum_{j=1}^{N} S_j^k}\right|. \quad (28)$$

In Fig. 9 and Fig. 10, we show the charging fairness M_c^k and mobility fairness M_m^k of executing different balancing decisions from 5am to 11pm, respectively. As the fairness metrics defined in (27) and (28), the charging fairness M_c^k is a negative sum of total absolute difference between the global and local charging supply-demand ratios at time k. The mobility fairness M_m^k has a similar definition. For both of them, larger values means higher fairness, which are preferable. And higher fairness indicates better balanced resource

allocation and similar service quality for customers among different regions and time frames in the city.

By using our DRO method, the average mobility fairness and charging fairness are improved by 15.78% and 34.51%, respectively, compared to the non-robust method, and improved by 10.45% and 30.92%, respectively, compared to the baseline robust method. Our DRO method outperforms the non-robust method because it considers both supply and demand uncertainties when making EV balancing decisions. The DRO problem formulation, uncertainty set construction based on data and equivalent convex optimization form derivation procedures are all designed carefully to solve the challenge in a computationally tractable way. In contrast, the baseline methods in the literature either do not consider model uncertainties such as the non-robust methods, or only consider one type of uncertainty such as the robust optimization method with the demand uncertainty.

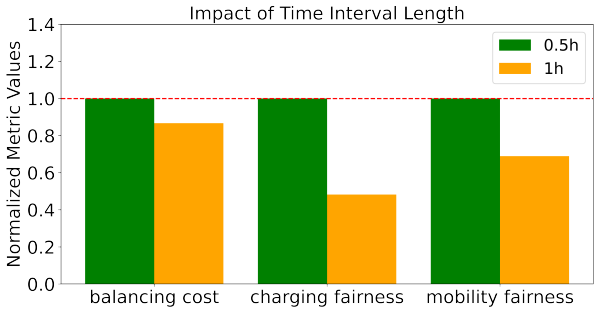


Fig. 11. With a shorter time interval length, the DRO solution results a higher total balancing cost in the same time frame, but better mobility and charging fairness.

In Fig. 11, we show the effect of different time interval lengths on our DRO method. We compare the total balancing cost, the mobility and charging fairness when using 0.5 hour and 1 hour as the time interval length, respectively. The metric values of the balancing cost, the mobility and charging fairness shown in Fig. 11 are normalized within $[0, 1]$ such that the highest result is fixed at value 1. With a shorter time interval l0,1ength, the DRO solution results a higher total balancing cost (a larger value in Fig. 11) in the same time frame, but better mobility and charging fairness. One interpretation is that there exists trade-off between different objectives for balancing EVs in the AMoD system. For instance, the algorithm can sacrifice balancing cost to get higher charging and mobility fairness by balancing the vehicles more frequently.

Remark (Practical implementation to address ITS open problems)**.** *Our proposed DRO method can be applied to the control of EV AMoD systems such as EV balancing and charge scheduling problems. Vehicle balancing is an important objective of EV AMoD systems [5], [38], [52]. By repositioning customer-free EVs, it aims to minimize the imbalance of the EV distribution caused by asymmetrical transportation demand [36]. Our method achieves lower balancing cost and higher mobility and charging fairness, thus contributing to congestion mitigation and transportation efficiency. Further, it is verified by using real-world data thus able to provide insights and helps autonomous taxis operation companies in their decisions.*

Remark (Limitations)**.** *(i) In this paper, we did not consider the potential impacts of political policies such as tax and subsidy in EV AMoD systems. For example, the EV promotion in Shenzhen is strongly supported by the government [58]. However, some literature has found that though tax and subsidy affect EV fleet evolution, there are little impacts on the EVs mobility and charging patterns [52]. Therefore, political factors are not emphasized in this work. (ii) The potential impacts of private EVs in AMoD systems are not considered. For instance, private EVs may share charging stations with autonomous EVs. However, according to some field studies in Shenzhen [38], few private EVs prefer utilizing fast charging stations because they have no need to leverage fast charging like commercial EVs for keeping normal business activities. Therefore, we did not explicitly include the impact of private EVs in AMoD balancing decisions. We do update the status and available spots of charging stations before making EV charging decisions in the designed DRO method, to mitigate the effects of other EVs that share the charging resource with the EV AMoD system. (iii) Our model shows good performance on the six-week data, and we will try to test it on datasets of longer time duration after we can have access to them.*

VII. Conclusion

Autonomous mobility-on-demand systems can provide efficient transportation services. However, with an increasing number of EVs, it is still challenging to improve the efficiency of AMoD systems under EVs supply uncertainty, due to limited charging facilities in cities and complicated charging dynamics. In this paper, we design a data-driven distributionally robust EV balancing method to minimize the worst-case expected cost under uncertainties of both passenger mobility demand and EV supply. In addition to reducing total vehicle balancing costs, we also balance the mobility and charging supply-demand ratios of different regions in the city. We propose efficient algorithms to construct distributionally uncertainty sets of the predicted mobility demand and EV supply. Then we derive an equivalent computationally tractable form of the distributionally robust EV balancing problem under the ellipsoid uncertainty sets constructed from historical data. Evaluations based on real-world E-taxi data show that the average total balancing cost is reduced by 14.49%, and the average passenger mobility fairness and EV charging fairness are improved by 15.78% and 34.51%, respectively. In the future, we will further evaluate our algorithm based on large-scale data from several years from multiple cities.

VIII. Appendix

A. Proof of Lemma 1

Proof. From the real data, we have historical $S^{1:\tau}$ and $r^{1:\tau}$: $\hat{S}^{1:\tau} = \{\hat{S}_1^1, ..., \hat{S}_N^\tau\}$, $\hat{r}^{1:\tau} = \{\hat{r}_1^1, ..., \hat{r}_N^\tau\}$. Then $l_g^k = \min\{\hat{S}_i^k/\hat{r}_i^k\}_{i\in\{1,...,N\}}$ is a global lower bound and $h_g^k = \max\{\hat{S}_i^k/\hat{r}_i^k\}_{i\in\{1,...,N\}}$ is a global upper bound in time k when there is no optimization on balancing. We call these two bounds global bounds. Let $l^k = l_g^k$, $h^k = h_g^k$ for all k, the supply-demand ratio after balancing should be contained

in this range. We denote these global lower and upper bounds of the mobility supply-demand ratio as $\bar{l}^k$ and $\bar{h}^k$.

Consider the optimization problem (29) that is similar to problem (2) except that problem (29) uses the global bounds as lower and upper bounds in the quality constraints.

$$\begin{aligned} \min_{\substack{X^{1:\tau},Y^{1:\tau},S^{1:\tau},D^{1:\tau},\\ U^{1:\tau},V^{2:\tau},O^{2:\tau},L^{2:\tau}}} \max_{\{F_r\in\mathcal{F}_r,F_c\in\mathcal{F}_c\}} & \mathbb{E}\left[J_D+\theta J_E\right] \\ \text{s.t. (4); (5); (6);} & \\ r_i^k-\bar{l}^k S_i^k-(D_i^k)^2=0, & \\ r_i^k-\bar{h}^k S_i^k+(U_i^k)^2=0, & \\ i\in\{1,\ldots,N\},\ k\in\{1,\ldots,\tau\}. & \end{aligned} \tag{29}$$

Suppose problem (29) has feasible optimal solutions $S^{*1:\tau}=\{S_1^{*1},...,S_N^{*\tau}\}$ and from the real data we have historical $\hat{r}^{1:\tau}=\{\hat{r}_1^1,...,\hat{r}_N^\tau\}$. Then we can compute l^k and h^k as following:

$$\begin{aligned} l^k&=\max\{l_g^k,\min\{S_i^{*k}/\hat{r}_i^k\}_{i\in\{1,\ldots,N\}}\},\\ h^k&=\min\{h_g^k,\max\{S_i^{*k}/\hat{r}_i^k\}_{i\in\{1,\ldots,N\}}\}. \end{aligned} \tag{30}$$

It is obvious that $l^k\geq l_g^k$ and $h^k\leq h_g^k$, so the range $[l^k,g^k]$ is no wider than $[l_g^k,h_g^k]$. We find feasible lower and upper bounds that no loose than the global bounds. □

B. Proof of Lemma 2

Proof. For the generalized EV balancing optimization problem, let's first check the modified constraints and objective functions. The modified constraints (14), (15) and (16) are still linear equality or inequality. So if the objective function in problem (19) is convex over all $S_{e,i}^k,V_{e,i}^k,O_{e,i}^k,L_{e,i}^k$, the modified objective function is a convex problem of the decision variables since the composition of affine or linear operation preserves convexity [[59], Chapter 3.2.2]. Now we only need to check the minimization problem (31) part in the generalized optimization problem is convex (since the maximization part is over the uncertain demand and supply parameters, not affected by the new formulation of $S_{e,i}^k,V_{e,i}^k,O_{e,i}^k,L_{e,i}^k$).

$$\min_{\substack{X^{1:\tau},Y^{1:\tau},S^{1:\tau},D^{1:\tau},\\ U^{1:\tau},V^{2:\tau},O^{2:\tau},L^{2:\tau}}} \left[J'_D+\theta J'_E\right] \text{s.t.(14); (15); (16); (8).} \tag{31}$$

By the definition of $J'_D(X_{1:E}^{1:\tau},Y_{1:E}^{1:\tau})$ in (18), it is a linear function of X_e^k and Y_e^k, hence also a convex function of X_e^k and Y_e^k, $\forall k=1,...,\tau;\ e=1,...,E$ according to the definition of convex. And J'_E is a convex function of all decision variables for any fixed value of $c_e^k,\forall k=1,...,\tau;\ e=1,...,E$: this is because the function of power $\frac{1}{x^a}$ is convex on scalar $x>0$ when $a>0$ [[59], Chapter 3.1.5], so $\forall k=1,...,\tau$, $J_E^{i,e,k}=\frac{c_{e,i}^k}{(T_i^k)^a}$ is a convex function on $T_i^k>0$. For more detail, for $T_i^k=\sum_{e=1}^E T_{e,i}^k=\sum_{e=1}^E[\sum_{j=1}^N y_{e,ji}^k-\sum_{j=1}^N y_{e,ij}^k]$, with a matrix $B^i\in\mathbb{R}^{N\times N}$ that $B_{ji}^i=1$, $B_{ij}=-1$, $Tr[B^iY_e^k]=T_{e,i}^k$. Then $J_E^{i,l,k}=1/(\sum_e Tr[B^iY_e^k])^a$ is a composition of convex function $1/x^a$ with an affine mapping: trace of the multiplication of metrics B^i and Y_e^k. It's an operation that preserves convexity [[59], Chapter 3.2.2]. Finally, $\theta>0$, $J'_D+\theta J'_E$, $J'_E=\sum_{k=1}^\tau\sum_{e=1}^L\sum_{i\in\sigma}J_E^{i,e,k}$ are both weighted sums of convex function, an operation that preserves convexity [[59], Chapter 3.2.1]. Hence, the minimization problem (31) is a convex optimization problem. □

C. Proof of Lemma 3

Proof. Without loss of generality, we prove the case for constraint parameter γ_α. Then according to the definition of quantiles $q_{\eta/2}^\gamma$ and $q_{1-\eta/2}^\gamma$ we have: $1-\eta=P(q_{\eta/2}^\gamma\leq\frac{\tilde{\gamma}_\alpha-\hat{\gamma}_\alpha}{s_\gamma}\leq q_{1-\eta/2}^\gamma)$. Under the assumption that the distribution of $\frac{\tilde{\gamma}_\alpha-\hat{\gamma}_\alpha}{s_\gamma}$ is close to the distribution of $\frac{\hat{\gamma}_\alpha-\gamma_\alpha}{s_\gamma}$ [[60], Chapter 18.6.1.1.2], we have: $P(q_{\eta/2}^\gamma\leq\frac{\tilde{\gamma}_\alpha-\hat{\gamma}_\alpha}{s_\gamma}\leq q_{1-\eta/2}^\gamma)=P(q_{\eta/2}^\gamma\leq\frac{\hat{\gamma}_\alpha-\gamma_\alpha}{s_\gamma}\leq q_{1-\eta/2}^\gamma)=P(\hat{\gamma}_\alpha-s_\gamma q_{1-\eta/2}^\gamma\leq\gamma_\alpha\leq\hat{\gamma}_\alpha-s_\gamma q_{\eta/2}^\gamma)=1-\eta$. So the probability that γ_α is between $\hat{\gamma}_\alpha-s_\gamma q_{\eta/2}^\gamma$ and $\hat{\gamma}_{\alpha1}-s_\gamma q_{1-\eta/2}^\gamma$ equals to $1-\eta$. □

D. Proof of Theorem 1

Proof. We notice that the uncertainty parameters are involved in both objective functions and constraints. We first use $\max_{F_r\in\mathcal{F}_r}\mathbb{E}(r_i^k)$ substitute r_i^k, $\min_{F_c\in\mathcal{F}_c}\mathbb{E}(c_i^k)$ substitute c_i^k in constraints (4) and (8). It's valid for minimizing the worst case: any uncertain values of r_i^k and c_i^k should meet the relationship with other decision variables in constraints (4) and (8). No matter which probability distribution is selected as the specific distribution to attain the worst case, the simplest worst case for single value of r_i^k and c_i^k is the case that the demand is really large that attains the maximal possible demand value while the supply is very small that attains the minimal possible supply value. And $\mathbb{E}(r_i^k)(\mathbb{E}(c_i^k))$ is the probability-weighted average of all its possible values. Then we transfer all constraints into functional formats as below:

$$\begin{aligned} f_{D_i^k}&=\max_{F_r\in\mathcal{F}_r}\mathbb{E}(r_i^k)-l_i^kS_i^k-(D_i^k)^2=0,\\ f_{U_i^k}&=\max_{F_r\in\mathcal{F}_r}\mathbb{E}(r_i^k)-h_i^kS_i^k+(U_i^k)^2=0,\\ f_{S_i^k}&=-S_i^k+X_i^k+V_i^k=0,\quad k=1,\ldots,\tau,\\ f_{V_i^{k+1}}&=-V_i^{k+1}+\sum_{j=1}^N P_{vji}^kS_j^k+\sum_{j=1}^N Q_{vji}^kO_j^k\\ &\quad-\max_{F_c\in\mathcal{F}_c}\mathbb{E}(-c_i^k)=0,\\ f_{O_i^{k+1}}&=-O_i^{k+1}+\sum_{j=1}^N P_{oji}^kS_j^k+\sum_{j=1}^N Q_{oji}^kO_j^k=0,,\\ f_{L_i^{k+1}}&=-L_i^{k+1}+Y_i^k+\sum_{j=1}^N P_{lji}^kS_j^k=0,L_i^k>0,\\ &\quad k=1,\ldots,\tau-1,\quad S_i^k>0,\quad k=1,\ldots,\tau. \end{aligned} \tag{32}$$

Let $f_D=[f_{D_1^1},f_{D_1^2},...,f_{D_1^\tau},...,f_{D_N^\tau}]^T\in\mathbb{R}^{N\tau}$ be a constraint function vector, for $i=1,\ldots,N,k=1,\ldots,\tau-1$, and f_U,f_S,f_V,f_O,f_L have the same definition but for computational convenient, if one's dimension is less than $N\tau$, we add 0 in corresponding missing positions to complete its dimension. For the primal maximization problem

$$\max_{F_r\in\mathcal{F}_r,F_c\in\mathcal{F}_c}\ \mathbb{E}\left[J_D+\theta J_E\right],\quad \text{s.t. (32)}$$

The primal objective function only contains uncertainty parameter c and is concave over c, because it's a linear function of c given other decision variables. The constraints are also all linear in r and c. For this primal problem, strong duality and Slater's theorem hold according to [59]. Then the Lagrange dual problem (33) can obtain its best upper bound:

$$\begin{aligned} &\min_{\lambda\succeq 0}\max_{F_r\in\mathcal{F}_r,F_c\in\mathcal{F}_c} J_{dual},\\ &J_{dual}=\mathbb{E}\left[J_D+\theta J_E\right]-(\lambda_U^T f_U+\lambda_S^T f_S+\lambda_V^T f_V\\ &\quad+\lambda_O^T f_O+\lambda_L^T f_L+\lambda_s^T S^{1:\tau}+\lambda_l^T L^{2:\tau}) \end{aligned} \tag{33}$$

Here $\lambda_U^T,\lambda_S^T,\lambda_V^T,\lambda_O^T,\lambda_L^T,\lambda_s^T,\lambda_l^T$ are the corresponding Lagrange multipliers and λ is defined as a concatenated vector combined by all these Lagrange multipliers. We have $\frac{c_i^k}{(T_i^k)^a}\geqslant 0$ and $c_i^k\geqslant 0$ by the definitions of J_E in (11), then for any vector $Z\in\mathbb{R}^{N\tau}$, $Z=[z_1^1,z_2^1,\ldots,z_1^\tau,z_2^\tau,\ldots,z_{N\tau}^\tau]^T$ that satisfies $0<\frac{1}{(T_i^k)^a}\leqslant z_i^k$, we also have

$$0\leqslant\sum_{k=1}^{\tau}\sum_{i=1}^{N}\frac{c_i^k}{(T_i^k)^a}\leqslant Z^Tc,$$

and the second inequality strictly holds when all $\frac{c_i^k}{(T_i^k)^a}=z_i^k$, for $i=1,\ldots,N$, $k=1,\ldots,\tau$. The constraints of problem (33) are independent of c, hence, for any c, the minmax problem (33) is equivalent to

$$\min_{\lambda\succeq 0}\max_{F_r\in\mathcal{F}_r,F_c\in\mathcal{F}_c} J'_{dual}\text{s.t.}\frac{1}{(T_i^k)^a}\leqslant z_i^k,\quad Z\in\mathbb{R}^{N\tau}, \tag{34}$$

where $J'_{dual}=\mathbb{E}\left[J_D+\theta Z^Tc\right]-(\lambda_U^Tf_U+\lambda_S^Tf_S+\lambda_V^Tf_V+\lambda_O^Tf_O+\lambda_L^Tf_L+\lambda_s^TS^{1:\tau}+\lambda_l^TL^{2:\tau})$. We separate J'_{dual} into three parts: $H_r=-(\lambda_U^T+\lambda_D^T)r, H_c=\theta J_E-\lambda_V^Tc, H_o=J'_{dual}-\mathbb{E}[H_c+H_r]$, only H_r contains all r, H_c contains all c. H_o can be put as a deterministic value given other decision variables. So we have the following maximization problem

$$\max_{r\sim F_r,c\sim F_c,F_r\in\mathcal{F}_r,F_c\in\mathcal{F}_c}\mathbb{E}[H_r+H_c]. \tag{35}$$

Since r and c are independent, problem (35) is equivalent to the separated maximization problem

$$\max_{r\sim F_r,F_r\in\mathcal{F}_r}\mathbb{E}[H_r]+\max_{c\sim F_c,F_c\in\mathcal{F}_c}\mathbb{E}[H_c]. \tag{36}$$

Problem (36) satisfies the conditions of Lemma 1 in [25], and the maximum expectation value of H_r+H_c for any possible $r\sim F_r, c\sim F_c$ where $F_r\in\mathcal{F}_r, F_c\in\mathcal{F}_c$ equals the optimal value of the problem

$$\begin{aligned} \min_{\substack{Q_r,q_r,v_r,t_r;\\ Q_c,q_c,v_c,t_c}}\quad & v_r+t_r+v_c+t_c\\ \text{s.t.}\quad & v_r\geqslant H_r-r^TQ_rr-r^Tq_r, Q_r,Q_c\succeq 0\\ & t_r\geqslant(\hat{\gamma}_r\hat{\Sigma}_r+\hat{r}\hat{r}^T)\cdot Q_r+\hat{r}^Tq_r\\ &\qquad+\sqrt{\hat{\omega}_r}\|\hat{\Sigma}_r^{1/2}(q_r+2Q_r\hat{r})\|_2,\\ & v_c\geqslant H_c-c^TQ_cc-c^Tq_c,\\ & t_c\geqslant(\hat{\gamma}_c\hat{\Sigma}_c+\hat{c}\hat{c}^T)\cdot Q_c+\hat{c}^Tq_c\\ &\qquad+\sqrt{\hat{\omega}_c}\|\hat{\Sigma}_c^{1/2}(q_c+2Q_c\hat{c})\|_2. \end{aligned} \tag{37}$$

Note that the first and third constraints about v_r and v_c is equivalent to $v_r\geqslant f_r(r^*)$ and $v_c\geqslant f_c(c^*)$ where $f_r(r^*)(f_c(c^*))$ is the optimal value of the following problem

$$\begin{aligned} &\max_r\quad H_r-r^TQ_rr-r^Tq_r\quad\text{s.t.}\quad r\geqslant 0.\\ &\max_c\quad H_c-c^TQ_cc-c^Tq_c\quad\text{s.t.}\quad c\geqslant 0. \end{aligned} \tag{38}$$

Since Q_r and Q_c are positive semi-defined, $H_r(H_c)$ is a linear function over $r(c)$, problem (38) is convex. Solving this problem by taking partial derivative over $r(c)$ without constraints, we have:

$$\begin{aligned} v_r&\geqslant\frac{1}{4}(q_r+\lambda_U+\lambda_D)^TQ_r^{-1}(q_r+\lambda_U+\lambda_D)\\ v_c&\geqslant\frac{1}{4}(q_c+\lambda_V-Z)^TQ_c^{-1}(q_r+\lambda_V-Z) \end{aligned} \tag{39}$$

By Schur complement, the above constraints are

$$\begin{aligned} &\begin{bmatrix} v_r & \frac{1}{2}(q_r+\lambda_U+\lambda_D)^T\\ \frac{1}{2}(q_r+\lambda_U+\lambda_D) & Q_r\end{bmatrix}\succeq 0,\\ &\begin{bmatrix} v_c & \frac{1}{2}(q_c+\lambda_V-Z)^T\\ \frac{1}{2}(q_c+\lambda_V-Z) & Q_c\end{bmatrix}\succeq 0, \end{aligned} \tag{40}$$

But these constraints are under the conditions of no constraints in problem (38). We still have another two constraints for $v_r(v_c)$ that $v_r\geqslant f_r(r=0)=0$ and $v_c\geqslant 0$. We then use the fact that min-min operations can be performed jointly and combine all constraints to reformulate problem (33) as (1). □

E. Non-robust Method Used in Experiment

The non-robust method treats the predictions of demand and supply as deterministic vales. Therefore, the non-robust method is solving a minimization problem (41) instead of a minimax problem.

$$\min_{\substack{X^{1:\tau},Y^{1:\tau},S^{1:\tau},D^{1:\tau},\\ U^{1:\tau},V^{2:\tau},O^{2:\tau},L^{2:\tau}}}\mathbb{E}\left[J_D+\theta J_E\right]\text{s.t. (4), (5), (6), (8),} \tag{41}$$

In the minimization problem (41), J_D is defined in Eq. (3) and J_E is defined in Eq. (11), θ is a positive coefficient which is chosen as the same value of the weight in the distributionally robust optimization EV balancing problem formulation (the robust minimax problem), i.e. Equation (12). We have discussed or proved that the objective functions J_D and J_E are both convex over all variables in the text. And all constraints (4), (5), (6), and (8) are convex (linear or quadratic). Therefore, this non-robust minimization problem can be solved through any convex optimization solver. This non-robust minimization problem is in a receding horizon control paradigm as well. So we are able to use the same parameters and real-sensing data used in our DRO method. The predictions of demand d and supply c are obtained from the same well-trained ARIMA model used in our DRO method.

REFERENCES

[1] R. Zhang, F. Rossi, and M. Pavone, "Model predictive control of autonomous mobility-on-demand systems," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1382–1389, IEEE, 2016.

[2] R. Zhang and M. Pavone, "Control of robotic mobility-on-demand systems: a queueing-theoretical perspective," *The International Journal of Robotics Research*, vol. 35, no. 1-3, pp. 186–203, 2016.

[3] W. J. Mitchell, C. E. Borroni-Bird, and L. D. Burns, *Reinventing the automobile: Personal urban mobility for the 21st century*. MIT press, 2010.

[4] G. Wang, X. Chen, F. Zhang, Y. Wang, and D. Zhang, "Experience: Understanding long-term evolving patterns of shared electric vehicle networks," in *The 25th Annual International Conference on Mobile Computing and Networking*, pp. 1–12, 2019.

[5] M. Ammous, S. Belakaria, S. Sorour, and A. Abdel-Rahim, "Optimal cloud-based routing with in-route charging of mobility-on-demand electric vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 7, pp. 2510–2522, 2018.

[6] Z. Hao, L. He, Z. Hu, and J. Jiang, "Robust vehicle pre-allocation with uncertain covariates," *Production and Operations Management*, vol. 29, no. 4, pp. 955–972, 2020.

[7] S. Chavhan, D. Gupta, B. Chandana, R. K. Chidambaram, A. Khanna, and J. J. Rodrigues, "A novel emergent intelligence technique for public transport vehicle allocation problem in a dynamic transportation system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 8, pp. 5389–5402, 2020.

[8] R. Iglesias, F. Rossi, R. Zhang, and M. Pavone, "A bcmp network approach to modeling and controlling autonomous mobility-on-demand systems," *The International Journal of Robotics Research*, vol. 38, no. 2-3, pp. 357–374, 2019.

[9] J. Zhen, D. Kuhn, and W. Wiesemann, "Mathematical foundations of robust and distributionally robust optimization," *arXiv preprint arXiv:2105.00760*, 2021.

[10] R. Iglesias, F. Rossi, K. Wang, D. Hallac, J. Leskovec, and M. Pavone, "Data-driven model predictive control of autonomous mobility-on-demand systems," in *2018 IEEE international conference on robotics and automation (ICRA)*, pp. 6019–6025, IEEE, 2018.

[11] F. Miao, S. Han, S. Lin, Q. Wang, J. A. Stankovic, A. Hendawi, D. Zhang, T. He, and G. J. Pappas, "Data-driven robust taxi dispatch under demand uncertainties," *IEEE Transactions on Control Systems Technology*, vol. 27, no. 1, pp. 175–191, 2017.

[12] Z. Tian, T. Jung, Y. Wang, F. Zhang, L. Tu, C. Xu, C. Tian, and X.-Y. Li, "Real-time charging station recommendation system for electric-vehicle taxis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 11, pp. 3098–3109, 2016.

[13] S. Wollenstein-Betech, M. Salazar, A. Houshmand, M. Pavone, I. C. Paschalidis, and C. G. Cassandras, "Routing and rebalancing intermodal autonomous mobility-on-demand systems in mixed traffic," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[14] F. Miao, S. Han, S. Lin, J. A. Stankovic, D. Zhang, S. Munir, H. Huang, T. He, and G. J. Pappas, "Taxi dispatch with real-time sensing data in metropolitan areas: A receding horizon control approach," *IEEE Transactions on Automation Science and Engineering*, vol. 13, no. 2, pp. 463–478, 2016.

[15] Z. Liu, J. Li, and K. Wu, "Context-aware taxi dispatching at city-scale using deep reinforcement learning," *IEEE Transactions on Intelligent Transportation Systems*, 2020.

[16] S. He, Y. Wang, S. Han, S. Zou, and F. Miao, "A robust and constrained multi-agent reinforcement learning framework for electric vehicle amod systems," *arXiv preprint arXiv:2209.08230*, 2022.

[17] H. Lin, X. Lin, H. Labiod, and L. Chen, "Toward multiple-phase mdp model for charging station recommendation," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[18] S. Faridimehr, S. Venkatachalam, and R. B. Chinnam, "A stochastic programming approach for electric vehicle charging network design," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 5, pp. 1870–1882, 2018.

[19] Y. Yuan, Y. Zhao, and S. Lin, "Sac: Solar-aware e-taxi fleet charging coordination under dynamic passenger mobility," in *2021 IEEE 60th Annual Conference on Decision and Control (CDC)*, IEEE, 2021.

[20] H. M. Abdullah, A. Gastli, and L. Ben-Brahim, "Reinforcement learning based ev charging management systems–a review," *IEEE Access*, vol. 9, pp. 41506–41531, 2021.

[21] A. Estandia, M. Schiffer, F. Rossi, J. Luke, E. C. Kara, R. Rajagopal, and M. Pavone, "On the interaction between autonomous mobility on demand systems and power distribution networks—an optimal power flow approach," *IEEE Transactions on Control of Network Systems*, vol. 8, no. 3, pp. 1163–1176, 2021.

[22] S. Zhao, X. Lin, and M. Chen, "Robust online algorithms for peak-minimizing ev charging under multistage uncertainty," *IEEE Transactions on Automatic Control*, vol. 62, no. 11, pp. 5739–5754, 2017.

[23] H. Yang, X. Xie, and A. V. Vasilakos, "Noncooperative and cooperative optimization of electric vehicle charging under demand uncertainty: A robust stackelberg game," *IEEE Transactions on vehicular technology*, vol. 65, no. 3, pp. 1043–1058, 2015.

[24] W. Wiesemann, D. Kuhn, and M. Sim, "Distributionally robust convex optimization," *Operations Research*, vol. 62, no. 6, pp. 1358–1376, 2014.

[25] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Operations research*, vol. 58, no. 3, pp. 595–612, 2010.

[26] V. Gabrel, C. Murat, and A. Thiele, "Recent advances in robust optimization: An overview," *European journal of operational research*, vol. 235, no. 3, pp. 471–483, 2014.

[27] F. Lin, X. Fang, and Z. Gao, "Distributionally robust optimization: A review on theory and applications," *Numerical Algebra, Control & Optimization*, vol. 12, no. 1, p. 159, 2022.

[28] Y. Wang, V. Puig, and G. Cembrano, "Set-membership approach and kalman observer based on zonotopes for discrete-time descriptor systems," *Automatica*, vol. 93, pp. 435–443, 2018.

[29] M. Pourasghar, V. Puig, and C. Ocampo-Martinez, "Comparison of set-membership and interval observer approaches for state estimation of uncertain systems," in *2016 European Control Conference (ECC)*, pp. 1111–1116, IEEE, 2016.

[30] F. Miao, S. He, L. Pepin, S. Han, A. Hendawi, M. E. Khalefa, J. A. Stankovic, and G. Pappas, "Data-driven distributionally robust optimization for vehicle balancing of mobility-on-demand systems," *ACM Transactions on Cyber-Physical Systems*, vol. 5, no. 2, pp. 1–27, 2021.

[31] R. Chen and C. G. Cassandras, "Optimal assignments in mobility-on-demand systems using event-driven receding horizon control," *IEEE Transactions on Intelligent Transportation Systems*, 2020.

[32] A. Mourad, J. Puchinger, and C. Chu, "A survey of models and algorithms for optimizing shared mobility," *Transportation Research Part B: Methodological*, vol. 123, pp. 323–346, 2019.

[33] J. Alonso-Mora, S. Samaranayake, A. Wallar, E. Frazzoli, and D. Rus, "On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment," *Proceedings of the National Academy of Sciences*, vol. 114, no. 3, pp. 462–467, 2017.

[34] X. Chen, F. Miao, G. J. Pappas, and V. Preciado, "Hierarchical data-driven vehicle dispatch and ride-sharing," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pp. 4458–4463, IEEE, 2017.

[35] F. Miao, S. Han, A. M. Hendawi, M. E. Khalefa, J. A. Stankovic, and G. J. Pappas, "Data-driven distributionally robust vehicle balancing using dynamic region partitions," in *Proceedings of the 8th International Conference on Cyber-Physical Systems*, pp. 261–271, 2017.

[36] G. Zardini, N. Lanzetti, M. Pavone, and E. Frazzoli, "Analysis and control of autonomous mobility-on-demand systems," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, pp. 633–658, 2022.

[37] J. Xu, R. Rahmatizadeh, L. Bölöni, and D. Turgut, "Taxi dispatch planning via demand and destination modeling," in *2018 IEEE 43rd Conference on Local Computer Networks (LCN)*, pp. 377–384, IEEE, 2018.

[38] G. Wang, W. Li, J. Zhang, Y. Ge, Z. Fu, F. Zhang, Y. Wang, and D. Zhang, "sharedcharging: Data-driven shared charging for large-scale heterogeneous electric vehicle fleets," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 3, pp. 1–25, 2019.

[39] F. Rossi, R. Iglesias, M. Alizadeh, and M. Pavone, "On the interaction between autonomous mobility-on-demand systems and the power network: Models and coordination algorithms," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 1, pp. 384–397, 2019.

[40] N. Noyan, G. Rudolf, and M. Lejeune, "Distributionally robust optimization with decision-dependent ambiguity set," *Optimization Online*, 2018.

[41] F. Luo and S. Mehrotra, "Distributionally robust optimization with decision dependent ambiguity sets," *Optimization Letters*, vol. 14, no. 8, pp. 2565–2594, 2020.

[42] S. He, L. Pepin, G. Wang, D. Zhang, and F. Miao, "Data-driven distributionally robust electric vehicle balancing for mobility-on-demand systems under demand and supply uncertainties," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2165–2172, IEEE, 2020.

[43] G. Wang, Y. Zhang, Z. Fang, S. Wang, F. Zhang, and D. Zhang, "Faircharge: A data-driven fairness-aware charging recommendation system for large-scale electric taxi fleets," *Proceedings of the ACM on*

Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 4, no. 1, pp. 1–25, 2020.

[44] L. Green, "Queueing theory and modeling," *Handbook of Healthcare Delivery Systems*, January 2011.

[45] D. Bertsimas, V. Gupta, and N. Kallus, "Data-driven robust optimization," *Mathematical Programming*, vol. 167, no. 2, pp. 235–292, 2018.

[46] P. M. Dixon, "Bootstrap resampling," *Encyclopedia of environmetrics*, vol. 1, 2006.

[47] R. J. Vanderbei *et al.*, *Linear programming*, vol. 3. Springer, 2015.

[48] A. Yurtsever, J. A. Tropp, O. Fercoq, M. Udell, and V. Cevher, "Scalable semidefinite programming," *SIAM Journal on Mathematics of Data Science*, vol. 3, no. 1, pp. 171–200, 2021.

[49] M. Grötschel, L. Lovász, and A. Schrijver, "The ellipsoid method and its consequences in combinatorial optimization," *Combinatorica*, vol. 1, no. 2, pp. 169–197, 1981.

[50] A. Agrawal, R. Verschueren, S. Diamond, and S. Boyd, "A rewriting system for convex optimization problems," *Journal of Control and Decision*, vol. 5, no. 1, pp. 42–60, 2018.

[51] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*. SIAM, 1994.

[52] G. Wang, F. Zhang, H. Sun, Y. Wang, and D. Zhang, "Understanding the long-term evolution of electric taxi networks: A longitudinal measurement study on mobility and charging patterns," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 11, no. 4, pp. 1–27, 2020.

[53] B. Du, H. Peng, S. Wang, M. Z. A. Bhuiyan, L. Wang, Q. Gong, L. Liu, and J. Li, "Deep irregular convolutional residual lstm for urban traffic passenger flows prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 3, pp. 972–985, 2019.

[54] O. Giraka and V. K. Selvaraj, "Short-term prediction of intersection turning volume using seasonal arima model," *Transportation Letters*, vol. 12, no. 7, pp. 483–490, 2020.

[55] M. Abbasi, A. Shahraki, and A. Taherkordi, "Deep learning for network traffic monitoring and analysis (ntma): A survey," *Computer Communications*, vol. 170, pp. 19–41, 2021.

[56] T. G. Dietterich and E. B. Kong, "Machine learning bias, statistical bias, and statistical variance of decision tree algorithms," tech. rep., Citeseer, 1995.

[57] Y. Yuan, D. Zhang, F. Miao, J. Chen, T. He, and S. Lin, "pˆ 2charging: Proactive partial charging for electric taxi systems," in *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)*, pp. 688–699, IEEE, 2019.

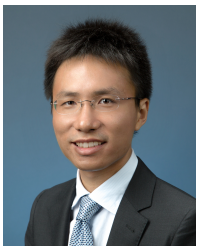
[58] M. Li, H. Ye, X. Liao, J. Ji, and X. Ma, "How shenzhen, china pioneered the widespread adoption of electric vehicles in a major city: Implications for global implementation," *Wiley Interdisciplinary Reviews: Energy and Environment*, vol. 9, no. 4, p. e373, 2020.

[59] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[60] R. Arnab, *Survey sampling theory and applications*. Academic Press, 2017.

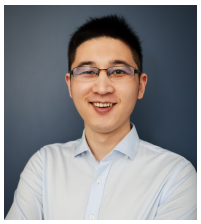
Sihong He is a PhD student in Computer Science and Engineering at University of Connecticut, Storrs, CT, USA. She received her bachelor's degree in Financial Mathematics from the Department of Mathematics, Southern University of Science and Technology and her master's degree in Statistics from the University of California, Irvine. Her current research interests include data-driven optimization, robust optimization, reinforcement learning and cyber-physical system, etc.

Zhili Zhang received B.Eng (2018) in computer science from University of Hong Kong and received M.Sc (2020) in computer science from New York University. He is currently working on his Ph.D in computer science and engineering at University of Connecticut, Storrs, CT. His research interests include cyber-physical system and reinforcement learning.



Shuo Han (S'08–M'14) received the B.E. and M.E. degrees in electronic engineering from Tsinghua University in 2003 and 2006, respectively, and the Ph.D. degree in electrical engineering from the California Institute of Technology in 2013. He was a Post-Doctoral Researcher with the Department of Electrical and Systems Engineering, University of Pennsylvania. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, University of Illinois at Chicago. His current research interests include developing rigorous frameworks for data-driven decision making that enable reliable and efficient operations of networked cyber-physical systems.

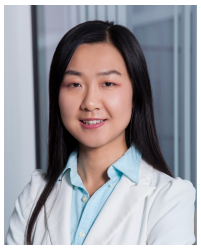
Lynn Pepin graduated from the University of Connecticut Computer Science and Engineering undergraduate program in 2018 and then joined as a PhD student under the advisory of Prof. Fei Miao. Since then, Lynn has worked on cybersecurity and machine learning research in the area of power systems and autonomous vehicles.



Guang Wang is an Assistant Professor at the Department of Computer Science, Florida State University. Before that, he was a Postdoctoral Research Associate at Massachusetts Institute of Technology. He obtained his Ph.D. degree in Computer Science, Rutgers University, New Brunswick, NJ, USA. He is interested in mobile computing, cyber-physical systems, big data analytics, and machine learning. His technical contributions have led to more than 40 publications in premium conferences and journals.

Desheng Zhang (M'11) is an Assistant Professor with the Department of Computer Science, Rutgers University. He is broadly concentrated on bridging cyber-physical systems and big urban data by technical integration of communication, computation, and control in data-intensive urban systems. He is focused on the life cycle of big data-driven urban systems, from multisource data collection to streaming-data processing, heterogeneous-data management, visualization, privacy, service design, and deployment in complex urban settings.

Jack Stankovic is the BP America Professor in the Computer Science Department at the University of Virginia, Charlottesville, Virginia, 22903, USA, and director of the Link Lab. His research interests include smart and connected health, cyberphysical systems, and the Internet of Things. Stankovic received a Ph.D. from Brown University. He is a Fellow of IEEE and the Association for Computing Machinery. Contact him at stankovic@virginia.edu.



Fei Miao (S'13-M'16) is an Assistant Professor in the Department of Computer Science and Engineering at the University of Connecticut. Previously, she was a Postdoc Researcher at the University of Pennsylvania. She received the Ph.D. degree in Electrical and Systems Engineering from the University of Pennsylvania in 2016. Her research interests include learning, optimization and control to assure safety, efficiency, robustness and security of cyber-physical systems. Dr. Miao is the recipient of the NSF CAREER Award, and Best Paper Award at the 12th ACM/IEEE International Conference on Cyber-Physical Systems.