Joint Order Dispatch and Charging for Electric Self-Driving Taxi Systems

1Guiyun Fan, 1Haiming Jin*, 1Yiran Zhao, 2Yiwen Song, 1Xiaoying Gan, 1Jiaxin Ding, 3Lu Su, 1Xinbing Wang
1Shanghai Jiao Tong University, Shanghai, China
2Carnegie Mellon University, Pittsburgh, USA, 3Purdue University, West Lafayette, USA
Email: 1{fgy726, jinhaiming, zhaoiyiran, ganxiaoying, jiaxinding, xwang8}@sjtu.edu.cn
2yiwens2@andrew.cmu.edu, 3lusu@purdue.edu

Abstract—Nowadays, the rapid development of self-driving technology and its fusion with the current vehicle electrification process has given rise to electric self-driving taxis (es-taxis). Foreseeably, es-taxis will become a major force that serves the massive urban mobility demands not far into the future. Though promising, it is still a fundamental unsolved problem of effectively deciding when and where a city-scale fleet of es-taxis should be charged, so that enough es-taxis will be available whenever and wherever ride requests are submitted. Furthermore, charging decisions are far from isolated, but tightly coupled with the order dispatch process that matches orders with es-taxis. Therefore, in this paper, we investigate the problem of joint order dispatch and charging in es-taxi systems, with the objective of maximizing the ride-hailing platform’s long-term cumulative profit. Technically, such problem is challenging in a myriad of aspects, such as long-term profit maximization, partial statistical information on future orders, etc. We address the various arising challenges by meticulously integrating a series of methods, including distributionally robust optimization, primal-dual transformation, and second order conic programming to yield far-sighted decisions. Finally, we validate the effectiveness of our proposed methods though extensive experiments based on two large-scale real-world online ride-hailing order datasets.

I. INTRODUCTION

Recent years have witnessed the rapid development of autonomous vehicles (AVs), empowered by a variety of sensors (e.g., camera, Lidar, mmWave radar) to accurately sense and perceive their surroundings, as well as advanced AI and control algorithms to adaptively drive themselves without human involvement. Nowadays, years of autonomous vehicle research are on the verge of being transformed into a revenue-producing business. Waymo, one of the leading companies in the AV industry, has started to open its self-driving taxi service to the public in the Phoenix area with a fleet of about 400 AVs. On the other side of the world, Chinese companies, such as DiDi and Baidu, are also launching their self-driving taxi services in certain public roads of Shanghai and Beijing.

Another technological trend that almost coincides with vehicle automation is vehicle electrification. Due to the environment friendly nature, several countries such as China and the United States have extensively adopted electric taxis to replace traditional gasoline ones. Although the early versions of self-driving taxis are mostly hybrids to enable continuous on-board sensing and computation, researchers and industry predict that all AVs will be electric not far into the future. In fact, Waymo has already started to shift its current hybrid self-driving taxis to all-electric ones. As a result, we have every reason to envision that fleets of electric self-driving taxis (es-taxis) will eventually be among the major forces that serve the massive transportation demands of urban residents.

However, though promising, efficiently managing a city-scale es-taxi fleet is still a fundamental yet unsolved problem. Different from refueling a gasoline taxi, charging an electric taxi’s battery from empty to full could take thirty minutes to a few hours. Currently, when and where to charge the taxis, which are critical for the number of ride orders they could serve, are decided by the drivers themselves. However, when operating an es-taxi fleet, the ride-hailing platform has to take over the responsibility of making appropriate charging decisions. Furthermore, the charging decisions are far from isolated, but tightly coupled with the order dispatch process that matches ride orders with available es-taxis. That is, charging affects the pool of es-taxis to which the ride orders could be dispatched, whereas order dispatch implicitly changes the locations and battery levels of es-taxis which eventually influences the spatio-temporal charging demands within a city.

Motivated by the aforementioned facts, in this paper, we investigate the problem of joint order dispatch and charging in an urban electric self-driving taxi system, with the objective of maximizing the platform’s long-term cumulative profit. However, solving such problem is challenging in a myriad of aspects. Next, we would shed some light upon the philosophies behind how we address the various arising challenges.

To summarize, this paper makes the following contributions.

- To the best of our knowledge, this paper is the first work that systematically and jointly optimizes the order dispatch and charging policy for urban electric self-driving taxi systems, with the objective of maximizing the ride-hailing platform’s long-term cumulative profit.
- We propose a novel solution approach that meticulously integrates stochastic dynamic programming with distributionally robust optimization, a carefully designed linear decision rule, and a series of primal-dual transformations. The proposed methods achieve robust profit maximization...

* Corresponding author.
1https://techcrunch.com/2020/10/08.
in a computationally efficient manner with only partial statistical knowledge of future orders.

- We conduct extensive experiments on two large-scale real-world online ride-hailing order datasets, including the one collected in Chengdu, China, consisting of 4,333,133 orders, from November 1st to 30th, 2016, and the other one collected in Haikou, China, consisting of 12,006,775 orders, from May 1st to October 30th, 2017. Our experimental results based on the above datasets validate the effectiveness of the methods proposed in this paper.

II. RELATED WORK

In recent years, a considerable amount of research effort [1–8] has been devoted to the development of self-driving technology. Specifically, these works propose solutions for collaborative platooning [1], road map data exchange [2], intersection management [3], misbehavior detection [4], routing [5], content delivery [6–8], as well as many others. In contrast, this paper focuses on developing joint order dispatch and charging methods for es-taxi systems, which has not been studied by the above works.

As aforementioned, our effort in this paper focuses on two major aspects, namely order dispatch and charging, which have been well investigated in the context of ride-hailing systems [9–26], but mostly in an isolated manner. On one hand, [9–14] aim at designing algorithms that efficiently match orders with available vehicles via reinforcement learning [9–11] and combinatorial optimization [12–14] without taking into consideration the issue of charging. On the other hand, [15–27] focus on the scheduling and pricing of charging services, as well as the planning of charging facilities without considering the issue of order dispatch. Specifically, [15–22] propose scheduling algorithms to efficiently match dynamic charging demands with available chargers, [23] aim at determining the charging prices which could maximize the profit of the charging serving provider, and [24–26] develop planning algorithms to generate locations for charging facilities that could serve potential charging demands in a cost-efficient manner. Due to the fact that charging and order dispatch are two tightly coupled processes for es-taxi systems, we jointly carry out decision making in these two aspects, which is different from the above works.

Furthermore, there are also other highly related past literature [28–37], but essentially investigate problems that are orthogonal to the one we address in this paper, including determining appropriate prices for ridesharing trips [28–30], as well as routing [31–34] and fleet management [35–37] in ridesharing systems. Different from the aforementioned existing works, this paper investigates the joint order dispatch and charging problem for es-taxi systems, aiming to maximize the platform’s long-term cumulative profit.

III. PRELIMINARY

A. System Overview

We consider an urban electric self-driving taxi (UEST) system, where a cloud-based platform operates \( C \in \mathbb{Z}^+ \) electric self-driving taxis (es-taxis) within a city to serve customers’ traveling requests. The platform tracks the locations of the es-taxi by collecting their GPS coordinates every few seconds.

In the spatial dimension, the platform divides the entire city into \( N \) equal-size service regions\(^4\), denoted as \( \mathcal{N} = \{1, 2, \ldots, N\} \). In the temporal dimension, the platform makes the order dispatch and charging decisions periodically, but at rather different frequencies. Oftentimes in practice, a ride-hailing platform dispatches orders after every few seconds, whereas it is usually not necessary to charge an es-taxi after at least tens of minutes. As a result, we define the decision period of order dispatch as a time slot, and that of charging as a time window which lasts for \( h \in \mathbb{Z}^+ \) time slots. We consider a planning horizon (usually one day) that contains \( S \in \mathbb{Z}^+ \) time windows, and thus \( H = S \times h \) time slots. We then denote the set of time slots and time windows as \( \mathcal{H} = \{1, 2, \ldots, H\} \) and \( \mathcal{S} = \{1, 2, \ldots, S\} \), respectively.

Next, we introduce a third type of temporal granularity, referred to as charging interval, which contains \( m \in \mathbb{Z}^+ \) time slots and represents the unit length of time an es-taxi charges its battery. We let \( G \in \mathbb{Z}^+ \) denote the maximum number of charging intervals required to fully charge an es-taxi, and let \( \mathcal{G} = \{1, 2, \ldots, G\} \). Note that our model allows \( h \) and \( m \) to be flexibly set according to real-world parameters. Figure 1 gives an example of the partition of the planning horizon with \( h = 12 \) and \( m = 7 \).

![Figure 1: An example of the partition of the planning horizon, where a time window contains 12 time slots and a charging interval contains 7 time slots, i.e., \( h = 12 \) and \( m = 7 \).](image)

In our model, we consider uniformly discretized electricity levels, denoted as \( \mathcal{K} = \{1, 2, \ldots, K\} \), where \( K \) represents the maximum electricity level of an es-taxi. Next, we define the type of an es-taxi and an order in the Definitions 1 and 2.

**Definition 1** (Type \( k \) Es-Taxi). An es-taxi with electricity level \( k \in \mathcal{K} \) is called a type \( k \) es-taxi.

Note that as the electricity level of an es-taxi usually changes over time, its type changes accordingly.

**Definition 2** (Type \( k \) Order). An \((i, j)\) order in our UEST system refers to a request submitted by a customer to travel from region \( i \) to region \( j \). We use \( e_{ij} \) to denote the level of electricity consumption of an \((i, j)\) order, and such order is a type \( k \) order, if \( e_{ij} = k \).

Note that if a type \( k \) es-taxi in region \( i \) is dispatched to serve an \((i, j)\) order, it will belong to type \( k - e_{ij} \) after arriving at

\(^4\)We consider service region as the smallest unit for location, and do not distinguish the locations within a service region. Thus, in our experiments in Section VI, we reasonably set each service region as a hexagon area with a side length of 1km.
As time goes by in time slot $t$, customers arrive at the UEST system and submit their orders to the platform.

- At the end of time slot $t$, the platform observes (i) the number of $(i,j)$ orders in the system that have not yet been dispatched, denoted as $O_{ijt}$, for each region pair $(i,j)$, (ii) the number of available es-taxis that belong to each type $k$ in each region $i$, denoted as $x_{ijt}^k$, and (iii) the number of available chargers in each region $i$, denoted as $a_{ijt}$.

- Based on the above observations, the platform makes the order dispatch decision $\mathbf{y}$ by deciding the number of each type $k$ es-taxis allocated to serve $(i,j)$ orders for each type $k$ and region pair $(i,j)$, denoted as $w_{ijt}^k$.

- Finally, available es-taxis either pick up costumers, or charge their batteries at the nearest charging station, or wait to be dispatched in the future time slots according to the decisions of the platform.

We let $\beta$ denote the number of electricity levels an es-taxi could charge in one charging interval. Then, at the end of time slot $t$, if the platform decides to charge a type $k$ es-taxi for $g$ charging intervals, it will belong to type $k + \beta g$ at the end of time slot $t + gm$.

### B. Problem Description

According to the workflow of our UVEST system, when the platform makes the order dispatch and charging decisions at the end of time slot $t$, it has the knowledge of the es-taxi vector $x_t = [x_{ijt}^k] \in \mathbb{R}^{N \times K}$, the order vector $O_t = [O_{ijt}] \in \mathbb{R}^{N \times T}$, and the charger vector $a_t = [a_{ijt}] \in \mathbb{R}^N$. The platform’s decisions at the current time slot naturally influence the number of available es-taxis steps into the future, and thus affect the system’s future order-serving ability. In order to fully capture such long-term effect, apart from the above observations, future orders should also be considered each time the platform makes a decision.

However, as the future orders have not yet arrived at the system in the current time slot, we treat $O_t$ in the future as a random matrix. We use $\mathbb{P}$ to denote the probability distribution of $O_{[H]} = (O'_{1}, O'_{2}, \cdots, O'_{H})'$. In practice, as it is usually rather difficult to obtain the precise knowledge of $\mathbb{P}$, our model only takes as input partial statistical knowledge about future orders, such as the first or second moment of $\mathbb{P}$.

Clearly, the platform receives revenue from order dispatch and incurs expense during charging. Next, we define the platform’s revenue and expense generated by the decisions in one time slot in Definition 3 and 4.

**Definition 3 (Revenue).** Let $p_{ijt}$ be the revenue from dispatching an $(i,j)$ order at the end of time slot $t$. We define the total revenue the platform receives from the order dispatch decision $\mathbf{w}_t = [w_{ijt}^k] \in \mathbb{R}^{N \times K}$ as $R_t(\mathbf{w}_t) = \sum_{i,j} \in \mathbb{P} p_{ijt} \left( \sum_{k \in K} w_{ijt}^k \right)$.

**Definition 4 (Expense).** Let $\theta_{gt}$ be the expense of charging for $g \in G$ charging intervals from the end of time slot $t$. The total expense of the platform incurred by the charging decision $c_t = [c_{igt}] \in \mathbb{R}^{NGK}$ is defined as $E_t(c_t) = \sum_{i \in N, k \in K, g \in G} \theta_{gt} c_{igt}$.

In this paper, we aim to maximize the long-term cumulative profit over the entire planning horizon defined in Definition 5.

**Definition 5 (Profit over Planning Horizon).** Given the charging decision $c_t$ and order dispatch decision $\mathbf{w}_t$, we define the platform’s immediate profit at the end of time slot $t$ as $f_t(c_t, \mathbf{w}_t) = R_t(\mathbf{w}_t) - E_t(c_t)$. Then, the cumulative profit of the platform over the whole planning horizon is defined as $F = \sum_{t \in H} f_t(c_t, \mathbf{w}_t)$.

In a nutshell, by jointly making the appropriate charging and order dispatch decisions at the end of each time slot $t$, we aim to maximize the cumulative profit of the platform over the whole planning horizon with partial statistical knowledge on future orders.

### IV. Mathematical Formulation

In this section, we formally formulate the problem which we aim to address in this paper. Mathematically, taking the randomness of future order arrivals into consideration, we formulate the platform’s decision-making problem at the end of each time slot $t$ as the following profit-maximizing (pm) stochastic dynamic program.

$$F_t(x_t, a_t, O_t) = \max_{(c_t, \mathbf{w}_t) \in A_t} \left\{ F_{t+1}(x_{t+1}, a_{t+1}, O_{t+1}) + E_t \left[ F_{t+1}(x_{t+1}, a_{t+1}, O_{t+1}) \right] \right\},$$

where

$$x_{(t+1)}^{k} = x_{it}^{k} + \sum_{j \in N} \left( w_{j(t+1)-j}^{k} - w_{ijt}^{k} \right)$$

$$+ \sum_{g \in G} \left( c_{igt}^{k} - c_{igt}^{k} \right), \forall k \in K, \forall i \in N$$

$$a_{(t+1)} = a_{it} + \sum_{k \in K, g \in G} \left( c_{igt}^{k} - c_{igt}^{k} \right), \forall i \in N$$

where $l_{j(t+1)}$ represents the number of time slots it takes for an es-taxi to travel from region $j$ to region $i$ at time slot $t$, and the terminal profit $F_{H+1}(x_{H+1}, a_{H+1}, O_{H+1}) = 0$.

Note that $\theta_{gt}$ depends on $t$ due to the time varying nature of the electricity price.

Note that $\theta_{gt}$ depends on $t$ due to the time varying nature of the electricity price.
Next, we explain in detail the objective function and the constraints of \( \text{pm} \).

Objective Function: As in Definition 5, the platform’s immediate profit is denoted as \( f_i(c_t, w_t) \). Then, we use \( \mathbb{E}_\rho[F_{t+1}(x_{t+1}, a_{t+1}, O_{t+1})] \) to represent its expected future profit. Thus, the Objective Function (1) of \( \text{pm} \) aims to maximize the platform’s cumulative profit from the end of time slot \( t \) to the end of the planning horizon.

Constraints:
- Constraint (2) ensures that, for each region \( i \) and type \( k \), the number of available es-taxis at the end of time slot \( t+1 \) equals to \( x^k_{it} \) plus the increment \( \sum_{j \in N} (w^k_{jiti}(t+1-t_i) - w^k_{ijt}) \) brought by the order dispatch decisions, and plus the increment incurred by the charging decisions, i.e., \( \sum_{g \in G} (c^k_{igt}(t+1-gm) - c^k_{igt}) \).
- Constraint (3) defines the relationship between the number of available chargers at the ends of time slots \( t+1 \) and \( t \). That is, at each region \( i \), \( a_{it} \) is obtained by adding to \( a_{it} \) the number of chargers released at the end of time slot \( t+1 \), represented by \( \sum_{k \in K, g \in G} (c^k_{igt}(t+1-gm)) \), and subtracting the number of chargers allocated to es-taxis by the charging decisions at the end of time slot \( t \), represented by \( \sum_{k \in K, g \in G} c^k_{igt} \).
- The decision variables \( (c_t, w_t) \) of \( \text{pm} \) at the end of time slot \( t \) belong to a feasible region \( A_t \), which is formally defined in the following Definition 6.

**Definition 6 (Feasible Region of Decision Variables).** The feasible region \( A_t \) of the decision variables \( (c_t, w_t) \) is defined as a set of possible values of \( (c_t, w_t) \) that satisfy the following Constraints (4)-(8). That is, for each \( i \in N \),

\[
\begin{align*}
\sum_{g \in G} c^k_{igt} + \sum_{j \in N} w^k_{ijt} & \leq x^k_{it}, & \forall k \in K, \quad (4) \\
\sum_{k \in K, g \in G} c^k_{igt} & \leq a_{it}, \quad (5) \\
\sum_{k \in K} w^k_{ijt} & \leq O_{ijt}, & \forall j \in N, \quad (6)
\end{align*}
\]

for each \( i \in N \), \( k \in K \), \( g \in G \),

\[
\begin{cases}
c^k_{igt} = 0, \text{ if } h \not= t, \text{ or } k + \beta g > K, \text{ or } t + gm > H; \\
c^k_{igt} \geq 0, \text{ otherwise},
\end{cases}
\]

and for each \( i, j \in N \),

\[
\begin{cases}
w^k_{ijt} = 0, \text{ if } k < e_{ij}, \\
w^k_{ijt} \geq 0, \text{ otherwise},
\end{cases}
\]

Next, we explain in detail Constraints (4)-(8) given in Definition 6.
- Constraint (4) ensures that, for each region \( i \) and each type \( k \), the sum of the number of es-taxis that are decided to charge and serve orders \( \sum_{g \in G} c^k_{igt} + \sum_{j \in N} w^k_{ijt} \) does not exceed that of the available es-taxis \( x^k_{it} \).
- Constraint (5) upper bounds \( \sum_{k \in K, g \in G} c^k_{igt} \), i.e., the number of es-taxis that are decided to charge at each region \( i \), by \( a_{it} \), i.e., the number of available chargers at that region.
- Constraint (6) guarantees that the sum of the es-taxis of all types allocated to \( (i, j) \) order \( \sum_{k \in K} w^k_{ijt} \) does not exceed the number of \( (i, j) \) orders.
- Constraints (7) denotes that the charging decisions are only made, when the end of a time slot \( t \) is also the beginning of a time window (i.e., \( b \not= t \)), the electricity level after charging does not exceed the highest possible level \( K \), and the time when charging completes does not exceed the end of the final time slot \( H \).
- Constraint (8) ensures that an es-taxi is only allocated orders that it has enough electricity to serve.

Thus far, we have finished formulating the \( \text{pm} \) problem, which is the one we aim to solve for the rest of the paper. Note that if the distribution \( \mathbb{P} \) is precisely known \textit{a priori}, existing methods (e.g., Monte Carlo and variance reduction techniques [40]) can be used to approximate the optimal solution. However, as aforementioned, the perfect knowledge of \( \mathbb{P} \) is relatively difficult to obtain in practice. Therefore, we consider that the platform only has partial knowledge of the distribution \( \mathbb{P} \) and adopt a robust optimization-based solution approach, which is further introduced in the next section.

**V. Solution Approaches**

In this section, we elaborate on our solution approaches for addressing the \( \text{pm} \) problem. We first transform \( \text{pm} \) into a robust optimization program with only partial knowledge on the distribution of future orders. Then, we perform duality transformation and further design a linear decision rule to obtain an SOCP with polynomial number of constraints, which could be solved efficiently.

**A. Robust Optimization Transformation**

As the platform only knows partial information of the distribution \( \mathbb{P} \), it is impossible to directly solve the \( \text{pm} \) problem. Therefore, instead of maximizing the platform’s profit as given in Equation 1, we choose to maximum the platform’s worst profit over all possible \( \mathbb{P} \)’s by adopting a distributionally robust optimization approach.

In what follows, we formally transform \( \text{pm} \) into a robust profit maximization (rpm) problem as

\[
F_i(x_t, a_t, O_t) = \max_{c, x_t \in A_t} \left\{ f_i(c_t, w_t) + \inf_{\rho \in \mathcal{W}} \mathbb{E}_\rho\left[ F_{t+1}(x_{t+1}, a_{t+1}, O_{t+1}) \right] \right\},
\]

where \( \text{rpm} \) has the same set of constraints as \( \text{pm} \), and \( \mathcal{W} \) represents the set of all possible distribution \( \mathbb{P} \)’s.

In our model, the platform only knows the expectation, range, and upper bounds on the variance of \( O_{ij} \) instead of the perfect knowledge of it, which is fairly reasonable in practice. Given such limited information, \( \mathcal{W} \) forms an ambiguity set of \( \mathbb{P} \) defined in the following Definition 7. For the ease of exposition, we simplify \( O_{ij} \) as \( O \) for the rest of this paper.

**Definition 7 (Ambiguity Set).** We let \( \mathcal{P} \) denote the set of all measures over \( \mathbb{R}^{N \times N \times H} \). Then, given the support set of \( O \) as \( \{O, \overline{O}\} \), the expectation of \( O \) as \( \mu \), and the upper bound of \( O_{ij} \)’s variance as \( \sigma^2_{ij} \), the set \( \mathcal{W} \subseteq \mathcal{P} \) defined by
Constraints 10-12 contains all possible \( \mathbb{P} \)'s and is referred to as an ambiguity set of \( \mathbb{P} \):

\[
\begin{align*}
\mathbb{P}(O \in [\mathbb{O}, \overline{\mathbb{O}}]) &= 1, \\
\mathbb{E}_O[O] &= \mu, \\
\mathbb{E}_O[(O_{ij} - \mu_{ij})^2] &\leq \sigma^2_{ij}, \forall i, j \in N, t \in \mathcal{H}. 
\end{align*}
\]

The parameters that appear in Constraints (10)-(12) can be obtained from real-world historical order datasets.

Next, we transform the non-linear functions of random variables inside the expectation in Constraint (12) into linear ones, which could significantly simplify the analysis. According to [41], we introduce an auxiliary random variable \( v = [v_{ijt}] \in \mathbb{R}^{N^2H} \), where \( v_{ijt} = (O_{ijt} - \mu_{ijt})^2 \). Furthermore, let us set \( Q \) be the joint distribution of \( (O, v) \), and \( \mathcal{P}_0 \) represent the set of all measures over \( \mathbb{R}^{N^2H} \times \mathbb{R}^{N^2H} \). Then, the set \( Q \subseteq \mathcal{P}_0 \) defined by Constraints (13)-(15) contains all possible \( Q \)'s, and is referred to as a lifted ambiguity set.

\[
\begin{align*}
\mathbb{P}(O_{ij} \in E) &= 1, \\
\mathbb{E}_O[O] &= \mu, \\
\mathbb{E}_O[v_{ijt}] &\leq \sigma^2_{ij}, \forall i, j \in N, t \in \mathcal{H}.
\end{align*}
\]

In Equation (13), \( \mathcal{E} \) denotes the support set of \( Q \), which satisfies constraints \( O \leq O \leq \overline{O} \) and \( (O_{ijt} - \mu_{ijt})^2 \leq v_{ijt} \leq \sigma_{ijt}^2 \) with \( \sigma_{ijt} = \max\{ (O_{ijt} - \mu_{ijt})^2, (O_{ijt} - \mu_{ijt})^2 \}, \forall i, j \in N, t \in \mathcal{H} \).

In the following, we show that the lifted ambiguity set \( Q \) and the ambiguity set \( W \) satisfy the relationship as given in Lemma 1.

**Lemma 1.** For all \( Q \in \mathcal{Q} \), the set of marginal distributions of \( O \) under \( Q \) is equivalent to the ambiguity set \( W \).

As the proof of Lemma 1 can be directly obtained from Proposition 1 in [41], we omit it for conciseness. By Lemma 1, we convert \( \text{rpm} \) into the following stochastic dynamic program under the lifted ambiguity set, referred to as the \( \text{lrpm} \) problem.

\[
\begin{align*}
F_r(x_t, a_t, O_t) &= \max_{\{e_t, w_t\} \in \mathcal{A}_t} \left\{ f_i(e_t, w_t) + \inf_{O_{ij} \in \mathcal{E}} \mathbb{E}_{O} [F_{t+1}(x_{t+1}, a_{t+1}, O_{t+1})] \right\},
\end{align*}
\]

which has the same set of constraints as \( \text{rpm} \). In the following Section V-B, we carry out further transformations to solve \( \text{lrpm} \).

**B. Duality Transformation**

In order to solve \( \text{lrpm} \), we next perform a duality transformation and convert it into the optimization program given in Theorem 1, which we refer to as the dual \( \text{lrpm} \) (dirpm) problem. Note that for the rest of the paper, we treat \( w, c, x, a \) and \( O \) as the ratios of the total number of ex-taxi \( C \) so as to perform the duality transformation.

**Theorem 1.** Dirpm is equivalent to the following optimization program \( \text{dirpm} \)

\[
\begin{align*}
\max_{\lambda \geq 0, \eta, \eta'(e_t, w_t) \in \mathcal{A}_t, (i, j, t) \in [t+1, H]} f_i(e_t, w_t) + \lambda + \eta' \mu + \sum_{i,j \in N, t \in [t+1, H]} \alpha_{ij} \sigma_{ij}^2 \\
\text{s.t. Constraints (2)-(3),} \\
\forall (O, v) \in \mathcal{E}, i \in N, I \in [t+1, H],
\end{align*}
\]

Combining the worst-case profit after time slot \( t \) with the decisions at time slot \( t \), we have the formulation given in Theorem 1.

\[
\begin{align*}
\mathcal{L}(O_{ij} \in E) &= 1, \\
\mathbb{E}_O[O] &= \mu, \\
\mathbb{E}_O[v_{ijt}] &\leq \sigma^2_{ij}, \forall i, j \in N, t \in \mathcal{H}.
\end{align*}
\]

As the \( \text{dirpm} \) problem given in Theorem 1 involves search-

\[8\mathbb{R}^{r,d}\] denotes the set of all measurable mappings from \( \mathbb{R}^r \) to \( \mathbb{R}^d \).
ing the best $c_I(\cdot)$ and $w_I(\cdot)$, $\forall I \in [t+1, H]$, over all the possible measurable mappings in $\mathbb{R}^{2N_t^I,NGK}$ and $\mathbb{R}^{2N_t^I,NK}$ respectively, which makes $\text{dlrpm}$ computationally intractable. Therefore, in order to handle such computational intractability, we propose to approximate each $c_I(\cdot)$ and $w_I(\cdot)$ using the linear decision rule (LDR) described in the following Section V-C, which restricts each $c_I(\cdot)$ and $w_I(\cdot)$ to be linear mappings of $(\mathbf{O}, \mathbf{v})$.

### C. Solution under Linear Decision Rule

As aforementioned, now we turn to solving $\text{dlrpm}$ in Theorem 1 under the LDR. That is, given any $t \in \mathbb{N}$, for each $I \in [t+1, H]$, we represent $c_I(\cdot)$ and $w_I(\cdot)$ as

$$
\begin{align*}
&c_I(\mathbf{O}_I, \mathbf{v}_I) = c^0_I + \sum_{i, j \in \mathbb{N}_t^I} c^1_{ij} I_{ij} + \sum_{i, j \in \mathbb{N}_t^I} c^2_{ij} I_{ij}, \\
w_I(\mathbf{O}_I, \mathbf{v}_I) = w^0_I + \sum_{i, j \in \mathbb{N}_t^I} w^1_{ij} I_{ij} + \sum_{i, j \in \mathbb{N}_t^I} w^2_{ij} I_{ij},
\end{align*}
$$

(23)

where $c^0_I, c^1_{ij}, c^2_{ij} \in \mathbb{R}^{NGK}$ and $w^0_I, w^1_{ij}, w^2_{ij} \in \mathbb{R}^{NK}$ are the parameters that we aim to solve.

Due to the fact that Constraints (17)-(22) should be satisfied for each $(\mathbf{O}, \mathbf{v}) \in \mathcal{E}$, $\text{dlrpm}$ then has infinite number of constraints. As a result, even though we introduce the above LDR, $\text{dlrpm}$ is not yet directly solvable. To solve the $\text{dlrpm}$ problem, we continue our transformation by applying again the duality technique to transform $\text{dlrpm}$ into a solvable form with finite number of constraints. In the following Theorem 2, we begin with the conversion of Constraint (22) to elaborate on the technical details.

#### Theorem 2. (Constraint (22) under LDR is equivalent to the following Constraints (25)-(27).)

\[
\begin{align*}
\lambda &+ \sum_{I \in \mathbb{N}, I \in [t+1, H]} \left( \sum_{i, j \in \mathbb{N}_t^I} p_{ij} w_{ij}^1 - \sum_{i, j \in \mathbb{N}_t^I} \theta_{gL} i_{ij}^1 \right) \\
&\geq \frac{\mathbf{3}}{2} \mathbf{O} - \mathbf{O} + \frac{1}{2} \mathbf{1}^T \mathbf{1} - \sum_{u \in \mathbb{N}^2[H-t]} \rho_u b_u, \\
\mathbf{H} - \mathbf{O} &+ \frac{1}{2} \mathbf{1}^T \mathbf{1} - \sum_{u \in \mathbb{N}^2[H-t]} \rho_u b_u,
\end{align*}
\]

(25)

Thus far, we have proved in Theorem 2 that Constraint (22) is equivalent to a collection of finite number of second order conic (SOC) representable constraints. Similar to the conversion in Theorem 2, we can also transform Constraints (19)-(21) in Theorem 1 into finite SOC representable constraints. In the following, we give the transformed expressions of Constraints (19)-(21) in Theorems 3-5, respectively. As the proofs of Theorems 3-5 are quite similar to that of Theorem 2, we omit them for conciseness.

#### Theorem 3. (For each $i \in \mathbb{N}, I \in [t+1, H], k \in \mathbb{K}$, Constraint (19) under LDR is equivalent to the following Constraints (36)-(38).)

\[
\begin{align*}
x_{ij} &- \sum_{j \in \mathbb{N}_t^I} w_{ij}^k - \sum_{g \in \mathbb{G}_t^I} \sum_{e \in \mathbb{E}_t^I} \left( w^k_{ij} - \sum_{j \in \mathbb{N}_t^I} w^k_{j1} - \sum_{g \in \mathbb{G}_t^I} \sum_{e \in \mathbb{E}_t^I} \theta^k_{g} i_{gkt} \right) \\
&+ \sum_{e \in \mathbb{E}_t^I} \left( \sum_{j \in \mathbb{N}_t^I} w^k_{ij} - \sum_{g \in \mathbb{G}_t^I} \sum_{e \in \mathbb{E}_t^I} \theta^k_{g} i_{gkt} \right) \\
&\geq \sum_{e \in \mathbb{E}_t^I} \left( \sum_{j \in \mathbb{N}_t^I} w^k_{ij} - \sum_{g \in \mathbb{G}_t^I} \sum_{e \in \mathbb{E}_t^I} \theta^k_{g} i_{gkt} \right) \\
&\geq \sum_{e \in \mathbb{E}_t^I} \left( \sum_{j \in \mathbb{N}_t^I} w^k_{ij} - \sum_{g \in \mathbb{G}_t^I} \sum_{e \in \mathbb{E}_t^I} \theta^k_{g} i_{gkt} \right),
\end{align*}
\]

(36)

This completes the proof.
Theorem 4. For each \(i \in N, I \in [t+1, H] \), Constraint (20) under LDR is equivalent to the Constraints (39)-(41).

\[
\begin{align*}
\alpha_{ij} &= \sum_{k \in K} \sum_{g \in G} \frac{a_{kg}}{g} t_{ijg} + \sum_{k \in K} \sum_{l \in [t-1]} \left( \sum_{g : x=g \Leftrightarrow I} \frac{b_{kg}}{g} (t_{ijg+1} - t_{ijg}) \right) \\
&+ \sum_{k \in K} \sum_{l \in [t-1]} \left( \sum_{g : x=g \Leftrightarrow I} \frac{c_{kg}}{g} t_{ijg} \right) - \sum_{k \in G} \frac{d_{kg}}{g} t_{ijg} \\
&\geq \gamma_{ij} - \sigma_{ij} + \frac{1}{2} \lambda_{ij} - \sum_{u \in [N^2 \times (t-1)]} \phi_{iju} b_u, \quad (39)
\end{align*}
\]

\[
\begin{align*}
\left( X_{ij}, Y_{ij} \right) &= \left( \frac{\lambda_{ij} - \gamma_{ij}}{0}, 0 \right) + \sum_{u \in [N^2 \times (t-1)]} \left( E_u \phi_{iju} + \lambda_{ij} c_u \right), \quad (40)
\end{align*}
\]

\[
\begin{align*}
\| \phi_{iju} \|_2 &\leq \lambda_{iju}, \quad \forall u \in [N^2 \times (t-1)], \quad (41)
\end{align*}
\]

where \( E_u = \begin{bmatrix} e_u^{0'} & 0' \end{bmatrix} \), \( b_u = \begin{bmatrix} -\mu_u \\ \frac{1}{\delta_u} \end{bmatrix} \), \( c_u = \begin{bmatrix} 0 \\ \frac{1}{\delta_u} \end{bmatrix} \), and elements in \( X_{ij} \) and \( Y_{ij} \) have similar expressions as Equations (28) and (29).

Theorem 5. For each \(i, j \in N, I \in [t+1, H] \), Constraint (21) under LDR is equivalent to the Constraints (42)-(44).

\[
\begin{align*}
\pi_{ij} - \psi_{ij} \sigma_{ij} - \frac{1}{2} \hat{\psi}_{ij} &\leq \sum_{u \in [N^2 \times (t-1)]} \psi_{iju} b_u, \quad (42)
\end{align*}
\]

\[
\begin{align*}
\xi_{ij} &= \left( \frac{\delta_{ij} - \xi_{ij}}{0}, 0 \right) + \sum_{u \in [N^2 \times (t-1)]} \left( D_u \psi_{iju} + \delta_{ij} c_u \right), \quad (43)
\end{align*}
\]

\[
\begin{align*}
\| \psi_{iju} \|_2 &\leq \delta_{iju}, \quad \forall u \in [N^2 \times (t-1)], \quad (44)
\end{align*}
\]

where \( D_u = \begin{bmatrix} e_u^{0'} & 0' \end{bmatrix} \), \( b_u = \begin{bmatrix} -\mu_u \\ \frac{1}{\delta_u} \end{bmatrix} \), \( c_u = \begin{bmatrix} 0 \\ \frac{1}{\delta_u} \end{bmatrix} \), and elements in \( \xi_{ij} \) and \( \psi_{ij} \) have similar expressions as Equations (28) and (29).

Theorem 6. \( \text{drlpm} \) under LDR is equivalent to the following second order cone program (SOCP)

\[
\begin{align*}
\max_{\alpha \geq 0, \lambda, \eta} &\quad f_{\alpha}(c_t, w_t) + \lambda \mu + \sum_{i, j \in N, I \in [t+1, H]} \alpha_{ij} \sigma_{ij}^2 \quad \text{s.t. Constraints (2)-(3), (23)-(27), (36)-(44).}
\end{align*}
\]

Thus far, we have shown that \( \text{drlpm} \) under LDR is equivalent to the SOCP given in Theorem 6. Next, we show in the following Theorem 7 that the SOCP in Theorem 6 has finite number of decision variables and constraints.

Theorem 7. The SOCP given in Theorem 6 has \( O\left(N^3(H^3 - t^3)K + N^4(H - t)^2 K \right) \) decision variables and \( O\left(N^3(H - t)^2 \right) \) constraints.

Proof. The dimensions of \( \alpha, \lambda, \) and \( \eta \) in Theorem 1 are \( O\left(N^2(H - t)\right) \), \( 1 \), and \( O\left(N^2(H - t)\right) \), respectively. The dimensions of \( c_t \) and \( w_t \) are \( O\left(N(H - t)K\right) \) and \( O\left(N^2 K\right) \), respectively. \( \forall I \in [t+1, H] \), the dimensions of \( c_t(\cdot) \) and \( w_t(\cdot) \) are \( N^2 I \times NGK \) and \( N^2 I \times N^2 K \), respectively. Thus, considering all of the above dimensions, we have in total \( O\left(N^3(H^3 - t^3)K + N^4(H - t)^2 K \right) \) decision variables. Next, we analyze the number of constraints that the SOCP has. Constraints (23) and (24) correspond to \( O\left(N^2(H - t)\right) \) constraints, Constraints (25)-(27) in Theorem 2 correspond to \( O\left(N^2(H - t)\right) \) constraints, Constraints (36)-(38) in Theorem 3 correspond to \( O\left(N^3(H - t)^2 K\right) \) constraints, Constraints (39)-(41) in Theorem 5 correspond to \( O\left(N^3(H - t)^2 \right) \) constraints, and Constraints (42)-(44) in Theorem 6 correspond to \( O\left(N^4(H - t)^2\right) \). Thus, the SOCP has overall \( O\left(N^3(H^3 - t^3)K + N^4(H - t)^2 K \right) \) constraints. □

Combining Theorems 6 and 7, we have finally transformed \( \text{rpm} \) into an SOCP with polynomial number of decision variables and constraints given in Theorem 6, which could be solved efficiently using commercial solvers (e.g. Gurobi and CPLEX). Therefore, we solve the SOCP in Theorem 6 at the end of each time slot \( t \) to obtain the charging and order dispatch decisions.

VI. PERFORMANCE EVALUATION

In this section, we introduce the baseline methods with which we compare our solution approach, the datasets, as well as the settings and results of our experiments.

A. Baseline Methods

As there is no existing literature that studies exactly the same problem as this paper, we compare our solution approach, referred to as long-term (L-T) in the rest of this paper, with the following three effective and meaningful baselines.

The first baseline is the greedy method, which is similar to the long-term approach except that the platform dispatches customers’ orders in current time slot as many as possible regardless of future order distributions. The second baseline is the force-charging (F-C) method, in which an es-taxi is forced to charge its battery to full, when its battery power is less than 10%. The third baseline is the myopic method, where the platform only maximizes the profit for two time windows.

B. Datasets and Experiment Settings

We conduct our extensive experiments with two large-scale public real-world datasets provided by the Didichuxing GAIA Initiative\(^9\). The first dataset contains the 4,333,133 online ride-hailing orders from November 1st to 30th, 2016, in Chengdu, China. The second dataset spans 6 months from May 1st to October 30th, 2017, with 12,006,775 online ride-hailing orders, in Haikou, China. An example of one piece of data from the Haikou dataset is given in the Table I. The Chengdu dataset has the same format as the Haikou dataset.

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pickup time</td>
<td>2017-05-19</td>
</tr>
<tr>
<td>Origin longitude</td>
<td>110.3665</td>
</tr>
<tr>
<td>Origin latitude</td>
<td>20.0059</td>
</tr>
<tr>
<td>Destination longitude</td>
<td>110.3645</td>
</tr>
<tr>
<td>Destination latitude</td>
<td>20.0353</td>
</tr>
</tbody>
</table>

The location and number of charging stations are obtained from a public website\(^10\) through the crawler we programmed. It shows that there are 448 stations with 813 chargers in Chengdu, China, and there are 154 stations with 369 chargers in Haikou, China.

We reasonably divide each city into 81 service regions, and set each service region as a hexagon area with a side length

---

\(^9\)https://outreach.didichuxing.com/research/opendata/en/
\(^10\)http://admin.bjev520.com/jsp/beiqi/pomap/do/index.jsp
of 1km. We map orders to the corresponding service regions based on the longitudes and latitudes of their origins and destinations. We extract the required parameters of the order distributions such as expectations from the actual datasets. We set the maximum cruising distance for an es-taxi as 300km, and the average speed of each es-taxi as 1km/min. The distance between two regions is set as the physical distance between two regions’ center points. Orders within 3km will be charged for 14RMB, and orders over 3km will be charged for 2RMB for each additional kilometer. In practice, it usually takes 1 hour to 2 hours to fully charge an es-taxi using a rapid charger. In our model, we thus reasonably set the length of each time slot to be 1 minute, the length of each charging interval to be \( m = 10 \), and it takes \( G = 10 \) charging intervals, i.e., 100minutes, to fully charge an es-taxi. The expense of charging for one charging interval to be \( \theta = 4 \). Based on our experiment results, the charging frequency of each es-taxi is 1.5 times per day on average. Furthermore, we set the other parameters according to Table II.

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Settings</th>
<th>( C )</th>
<th>( h )</th>
<th>( K )</th>
<th>( H )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chengdu</td>
<td>I</td>
<td>4000</td>
<td>12</td>
<td>300</td>
<td>[360, 720]</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>4000</td>
<td>12</td>
<td>300</td>
<td>720</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>[1500, 4500]</td>
<td>12</td>
<td>300</td>
<td>720</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>4000</td>
<td>12</td>
<td>[50, 350]</td>
<td>720</td>
</tr>
<tr>
<td>Haikou</td>
<td>V</td>
<td>2000</td>
<td>12</td>
<td>300</td>
<td>[360, 720]</td>
</tr>
<tr>
<td></td>
<td>VI</td>
<td>2000</td>
<td>12</td>
<td>300</td>
<td>720</td>
</tr>
<tr>
<td></td>
<td>VII</td>
<td>[500, 2300]</td>
<td>12</td>
<td>300</td>
<td>720</td>
</tr>
<tr>
<td></td>
<td>VIII</td>
<td>2000</td>
<td>12</td>
<td>[50, 350]</td>
<td>720</td>
</tr>
</tbody>
</table>

C. Experiment Results

Figures 2-9 demonstrate the comparison of the platform’s cumulative profit over the whole planning horizon yielded by our long-term method, and the baseline methods. Figures 2-5 are plotted under settings I-IV, and Figures 6-9 are plotted under settings V-VIII. We can easily observe from Figures 2-9 that the platform’s cumulative profits of our long-term method are far more than those of the baseline methods under all parameter settings.

Besides, we could observe in Figures 2 and 6 that the profit gaps between the long-term method and the other methods increase with the length of the planning horizon, which shows the superiority of the former approach in long-term decision making. Figures 3 and 7 show that the platform’s profit slightly reduces with a longer charging window. The reason is that a lower charging decision frequency may hold back more es-taxis that need to be charged, which reduces the number of available es-taxis for order dispatch. It can be observed from...
Figures 4 and 8 that platform’s profit increases with the total number of es-taxis, because more orders could be satisfied with more es-taxis. From Figures 5-9, we can observe that the platform’s profit increases with the maximum electricity level of an es-taxi, because the platform dispatches orders more accurately with a more fine-grained battery level division.

In order to illustrate the reasons behind the superiority of the long-term approach compared with the baseline methods, we further decompose the cumulative profit into the order dispatch revenue and the charging expense, and plot Figures 10-13. Figures 10 and 12 show that although the myopic solution has a lower charging expense, it serves far less orders than the other three approaches. In comparison, the long-term approach could serve the most orders while maintaining a relatively small charging overhead. Furthermore, the revenue gaps between the long-term method and the other approaches increase with the maximum number of time slots, which shows the power and necessity of adopting the former approach in real-world UEST systems where the platform makes decisions taking a farsighted view.

We can observe from Figures 11 and 13 that although the charging expense of the long-term method is close to that of the force-charging method, the former serves far more orders than the latter one. Such result indicates the charging decisions of our long-term method will guide es-taxis to charge at more reasonable time and duration than the baseline methods.

In Figures 14-17, we visualize the heat maps of the number of unserved orders generated by our experiments with $H = 720$ under setting I and setting V respectively, and show the results of four representative time slots and part of the regions in Chengdu. In these figures, a darker color represents a higher number of unserved orders. The darkest color indicates that more than 50% of the orders in this region are not served, and the lightest color represents that all orders in this region are served in the current time slot. The number of orders in each time slot at each region in Chengdu dataset varies from zero to 163. From these figures, we can easily observe that our long-term approach offers a much more satisfactory order serving than the other approaches.

**ACKNOWLEDGMENT**

This work was supported by NSF China (No. U21A20519, 62061146002, U20A20181, 61902244), and in part by Shanghai Municipal Science and Technology Commission Grant 19YF1424600.

**VII. CONCLUSION**

In this paper, we study the problem of joint order dispatch and charging in UEST systems, with the objective of maximizing the ride-hailing platform’s long-term cumulative profit. Such long-term optimization objective together with the partially known statistical knowledge on future orders makes the problem rather challenging to resolve. To address the various arising challenges, we propose a novel solution method that meticulously integrates stochastic dynamic programming with distributionally robust optimization, a carefully designed linear decision rule, a series of primal-dual transformations, and second order conic programming. In this way, our method achieves robust profit maximization in a computationally efficient manner with only partial statistical knowledge of future orders. To validate the effectiveness of our methods, we conduct extensive experiments on two large-scale real-world online ride-hailing order datasets. Our experiment results demonstrate that our method could guide es-taxis to charge at more reasonable time and durations, and thus could serve much more orders with a relatively small charging overhead compared with the baseline methods.