Joint Pricing and Rebalancing of Autonomous Mobility-on-Demand Systems

Salomón Wollenstein-Betech¹, Ioannis Ch. Paschalidis¹, and Christos G. Cassandras¹

Abstract—This paper studies optimal pricing and rebalancing policies for Autonomous Mobility-on-Demand (AMoD) systems. We take a macroscopic planning perspective to tackle a profit maximization problem while ensuring that the system is load-balanced. We begin by describing the system using a dynamic fluid model to show the existence and stability of an equilibrium (i.e., load balance) through pricing policies. We then develop an optimization framework that allows us to find optimal policies in terms of pricing and rebalancing. We first maximize profit by only using pricing policies, then incorporate rebalancing, and finally we consider whether the solution is found sequentially or jointly. We apply each approach on a data-driven case study using real taxi data from New York City. Depending on which benchmarking solution we use, the joint problem (i.e., pricing and rebalancing) increases profits by 7% to 40%.

Index Terms—Autonomous Mobility-on-Demand Systems; Revenue Maximization; Load balancing; Pricing; Optimization.

I. INTRODUCTION

With the rise of Mobility-on-Demand (MoD) services (e.g. Uber, Lyft, DiDi) and the rapid technological evolution of self-driving vehicles, we are closer to having Autonomous Mobility-on-Demand (AMoD) systems. A crucial step in the proper functioning of such a service is to define pricing, rebalancing and routing policies for the fleet of vehicles. This paper focuses on the first two issues, while the interested reader is directed to [1] for a discussion on routing and rebalancing.

Pricing policies play an important role as they modulate the inflow of customers traveling between regions in the network. As a result, the controller has the ability to choose prices such that the induced demand ensures a balanced load of customers and vehicles arriving at each location. Additionally, selecting prices enables the operator to modify demand such that the system can operate with smaller or larger fleet sizes. If we restrict a pricing policy to one that requires balancing the load in every node, we expect the solution to concentrate on balancing the network rather than choosing a set of prices to maximize profit. To give the pricing policy more flexibility, AMoD systems can leverage rebalancing policies, i.e., send empty vehicles from regions with excess supply of vehicles to regions with excess demand of trips (see Fig.1) with the objective of achieving higher profit.

†This work was supported in part by NSF under grants ECCS-1509084, DMS-1664644, CNS-1645681, IIS-1914792, and CMMI-1454737, by AFOSR under grant FA9550-19-1-0158, by ARPA-E’s NEXTCAR program under grant DEAR0000796, by the MathWorks, by the ONR under grant N00014-19-1-2571, and by the NIH under grant 1R01GM135930. We thank D. Sverdrin-Lisker for proofreading this paper.

¹ The authors are with the Division of Systems Engineering, Boston University, Brookline, MA 02446 USA {salomonw, cgc, yannisp}@bu.edu
methods in pricing, we consider the destination of a customer when designing the pricing policy. This allows the fleet controller to modulate demand in such a way that the system is balanced by solely adjusting prices. Additionally, we incorporate the rebalancing policy optimization framework in [13] and formulate a joint optimization model. We compare this joint strategy with four different methodologies. First by only finding optimal prices, second by only rebalancing the fleet, third by sequentially solving the rebalancing and then pricing of the system, and fourth by jointly estimating pricing and rebalancing with a unique surge price by origin. We apply each approach to two case studies; one, with simulated data; and another, with real taxi data from New York City.

**Organization:** The paper is organized as follows. In Section [II] we introduce the fluid model consisting of queues of customers and vehicles at every region. In Section [III] we show that the system is well-posed and establish the existence of a load balance equilibrium through the selection of prices. We also obtain local stability results. Next, in Section [IV] we state the problems of optimal pricing, optimal rebalancing and the joint formulation of these two. Then, we present case studies to assess the performance of the joint formulation in Section [V] Finally, in Section [VI] we conclude.

II. Model

In this section we present a steady-state deterministic fluid model to find optimal prices in an AMoD system while ensuring service to customers. This model is intended to serve as a relaxation of the corresponding stochastic queuing model where customers arrive according to a Poisson process and travel times are non-deterministic (usually exponentially distributed). The reason for making this relaxation is the flexibility it provides to perform analysis of the system.

Consider a fully-connected network $G = (N, A)$ where $N$ is the set of nodes (regions) $N = \{1, ..., N\}$ and $A = \{(i, j) : i, j \in N \times N\}$ is the set of arcs. A customer requests a ride in region $i$, receives a transportation service from the AMoD platform, and is charged a price composed of the product of a base and a surge price. The total price is $p_{ij} = p_{ij}^b u_{ij}$ where $p_{ij}^b, u_{ij}$ are the base and surge prices, respectively, for traveling from node $i$ to $j$. Throughout the paper, we will use the surge price $u_{ij}$ as our control variable, and we assume that $u_{ij} \geq 1$ as the platform is not willing to charge less than its base price.

We further assume that customers’ arrival rate is a function of the surge price, namely $\lambda_{ij}(u_{ij}) : \mathbb{R}_{\geq 0} \mapsto \mathbb{R}_{\geq 0}$ for a customer travelling from $i$ to $j$. This function is known as the williness-to-pay or the demand function. Let the base demand be $\lambda^b_{ij} = \lambda_{ij}(1)$, i.e., the demand rate of customers when the surge price is at its minimum.

As in [13], we use a queuing model for this system with two queues per region. We let $c_i(t) \in \mathbb{R}_{\geq 0}$ be the number of customers at region $i$ waiting to be assigned to a vehicle; and denote with $v_i(t) \in \mathbb{R}_{\geq 0}$ the number of available vehicles waiting in region $i$ at time $t$. Moreover, the AMoD provider assigns vehicles to customers located in the same region at a service rate $\mu_i$. We assume that $\mu_i > \sum_j \lambda^b_{ij}$, meaning that the platform assigns vehicles to customers faster than the rate at which customers arrive.

We assume that $\lambda_{ij}(\cdot)$ is monotonic decreasing for all $i, j \in N$, i.e., as price increases, the demand rate decreases.

**Assumption 1.** The function $\lambda_{ij}(\cdot)$ is monotonically decreasing $\forall i, j \in N$, i.e., as price increases, the demand rate decreases.

**Assumption 2.** There exists a surge price $u^{max}_{ij}$ for which $\lambda_{ij}(u^{max}_{ij}) = 0$, $\forall i, j \in N$.

A. Customer Dynamics

Consider a customer queue $c_i$ for each region $i \in N$ in the network. The queue dynamics are:

$$
\dot{c}_i = \begin{cases} 
\sum_j \lambda_{ij}(u_{ij}), & \text{if } v_i = 0, \\
0, & \text{if } v_i \geq 0 \text{ and } c_i = 0, \\
\sum_j \lambda_{ij}(u_{ij}) - \mu_i, & \text{if } v_i \geq 0 \text{ and } c_i \geq 0.
\end{cases}
$$

In order to express the customer dynamics with shorter notation we let $H(x)$ be an indicator function for positive values of $x$, and we use the following shorthand notation:

$\lambda_{ij} := \lambda_{ij}(u_{ij})$, $\lambda_i := \sum_j \lambda_{ij}$, $v_i := v_i(t)$, $c_i := c_i(t)$, $v^j_i := v_j(t - T_{ji})$, $c^j_i := c_j(t - T_{ji})$

where $\lambda_i$ is the total endogenous outgoing flow from node $i$; and $c^j_i$, $v^j_i$ are the rates of customer and vehicle arrivals at $t$ to region $i$ coming from $j$, respectively, i.e., $v^j_i$ is the rate of vehicles who departed from $j$ destined for $i$, $T_{ji}$ time units prior to the current time $t$. Then, we rewrite the customer dynamics in compact form can be written as follows:

$$
\dot{c}_i = \lambda_i(1 - H(v_i)) + (\lambda_i - \mu_i)H(c_i)H(v_i).
$$

Note that as a result of using a fluid model, the variables denoting the number of customers in a region are real-valued.
B. Vehicle Dynamics

The outflow rate corresponding to vehicles departing station $i$ is given by:

$$
\dot{v}_i^- = \begin{cases} 
-\lambda_i, & \text{if } v_i \geq 0 \text{ and } c_i = 0, \\
0, & \text{if } v_i = 0, \\
-\mu_i, & \text{if } v_i \geq 0 \text{ and } c_i \geq 0.
\end{cases}
$$

and more succinctly as

$$
\dot{v}_i^- = -\lambda_i H(v_i) + (\lambda_i - \mu_i) H(v_i) H(c_i). \tag{2}
$$

In addition, the rate at which customer-carrying vehicles arrive at station $i$ is given by:

$$
\dot{v}_i^+ = \sum_j (\lambda_j i H(v_j^i) - (\lambda_j - \mu_j) H(v_j^i) H(c_j^i)). \tag{3}
$$

Therefore, we can write the vehicle dynamics in compact form by adding (2) with (3) as

$$
\dot{v}_i = -\lambda_i H(v_i) + (\lambda_i - \mu_i) H(c_i) H(v_i) + \sum_j (\lambda_j i H(v_j^i) - (\lambda_j - \mu_j) H(v_j^i) H(c_j^i)). \tag{4}
$$

Hence the global system dynamics are expressed by the following differential equations

$$
\dot{c}_i = \lambda_i (1 - H(v_i)) + (\lambda_i - \mu_i) H(c_i) H(v_i), \tag{5a}
$$

$$
\dot{v}_i = -\lambda_i H(v_i) + (\lambda_i - \mu_i) H(c_i) H(v_i) + \sum_j (\lambda_j i H(v_j^i) - (\lambda_j - \mu_j) H(v_j^i) H(c_j^i)). \tag{5b}
$$

which describe a non-linear, time-delayed, time-invariant, right-hand discontinuous system.

III. WELL POSEDNESS, EQUILIBRIUM AND STABILITY

Similar to [13], we say that the system (5) is well posed if two conditions are satisfied: (i) for any initial condition, there exists a solution of the differential equations in (5), and (ii) the number of vehicles in the system remain invariant over time. In order to analyze the model, we use the framework of Filippov solutions [14]. Let us now give a proposition for the well-posedness of the system:

Proposition 1 (Well-posedness of fluid model).

1) For every initial condition in the fluid model represented in (5), there exist continuous functions $c_i(t) : \mathbb{R} \geq 0 \mapsto \mathbb{R} \geq 0$ and $v_i(t) : \mathbb{R} \geq 0 \mapsto \mathbb{R} \geq 0$, $i \in \mathcal{N}$, satisfying the system of equations in the Fillipov sense.

2) For all $t > 0$, the total number of vehicles is invariant and equal to $m = \sum_{i \in \mathcal{N}} v_i(0)$.

Proof: For the first claim, we use the framework developed by [15]. In particular, we check that all assumptions and conditions of [15, Thm II-1] are satisfied. This theorem, ensures the existence of Filippov solutions to the time-delayed differential equations with discontinuous right-hand sides.

To prove the second claim, we study the dynamics of the vehicles in the system, which we separate into two categories: vehicles who are in transit $v_i(t)$, and vehicles at a specific region $v_i(t)$. For the vehicles queued at $i$ we know their dynamics are as in (5b). For the vehicles in transit, we let the total be

$$
v_{ij}(t) = \int_{t-\tau_{ij}}^{t} \lambda_i H(v_i(\tau)) + (\lambda_i - \mu_i) H(c_i(\tau)) H(v_i(\tau)) \, d\tau,
$$

and their dynamics are

$$
\dot{v}_{ij}(t) = \lambda_i H(v_i) + (\lambda_i - \mu_i) H(c_i) H(v_i) - (\lambda_i H(v_i^i) + (\lambda_i - \mu_i) H(c_i^i) H(v_i^i)). 
$$

Moreover, the total number of vehicles in the system is

$$
m(t) = \sum_i v_i(t) + \sum_{ij} v_{ij}(t),
$$

with dynamics

$$
\dot{m}(t) = \sum_i \dot{v}_i(t) + \sum_{ij} \dot{v}_{ij}(t), \tag{6a}
$$

$$
= \sum_i \left( -\lambda_i H(v_i) + (\lambda_i - \mu_i) H(c_i) H(v_i) \right) \tag{6b}
$$

$$
+ \sum_{ij} \lambda_i H(v_j^i) - (\lambda_i - \mu_i) H(c_j^i) H(v_j^i) + \sum_{ij} \dot{v}_{ij},
$$

$$
= \sum_{ij} -\lambda_i H(v_i) + (\lambda_i - \mu_i) H(c_i) H(v_i) \tag{6c}
$$

$$
+ \sum_{ij} \lambda_j i H(v_j^i) - (\lambda_j - \mu_j) H(v_j^i) H(c_j^i) H(v_j^i) + \sum_{ij} \dot{v}_{ij}, \tag{6d}
$$

Note that to obtain the above result we have expanded the first sum term in (6a) using (5b), rearranged terms and found that $-\sum_i \dot{v}_i(t) = \sum_{ij} \dot{v}_{ij}(t) \implies \dot{m}(t) = 0$, which implies that the fleet size remains invariant over time.

A. Equilibria

We say that the system is in equilibrium if customer queues (and therefore, waiting times) do not grow infinite. We show the existence of an equilibrium in the fluid model (5) when we control the prices of every origin-destination pair. Additionally, we show that by having the ability to control the prices, one can have find multiple equilibria for a desired fleet size, giving the flexibility to AMoD managers to operate the system at different levels.

Theorem 1 (Existence of equilibria). Let $\mathcal{U}$ be a set of prices $u$, such that when $u \in \mathcal{U}$:

$$
\sum_j \lambda_{ij}(u_{ij}) - \lambda_{ji}(u_{ji}) = 0, \quad \forall i \in \mathcal{N},
$$

and let

$$
m_u := \sum_{ij} T_{ij} \lambda_{ij}(u_{ij}). \tag{8}
$$

Then, if $u \in \mathcal{U}$, and $m > m_u$, an equilibrium exists with $c = 0$ and $v > 0$. Otherwise no equilibrium exists.

Proof: Set $\dot{c}_i = 0$ and $\dot{v}_i = 0$ for all $i \in \mathcal{N}$. Then by using the customer system dynamics in (5a), we have:

$$
\lambda_i = \lambda_i H(v_i) - (\lambda_i - \mu_i) H(c_i) H(v_i), \tag{9}
$$

and since $\lambda_i < \mu_i$, the above equation just has a solution if $c_i = 0$ and $v_i > 0$ for all $i \in \mathcal{N}$. Setting $\dot{v}_i = 0$, and using the vehicle dynamics in (5b) we have

$$
0 = -\lambda_i H(v_i) + (\lambda_i - \mu_i) H(c_i) H(v_i) + \sum_j \lambda_{ji} H(v_j^i) - (\lambda_{ji} - \mu_j) H(c_j^i) H(v_j^i), \tag{10}
$$

where $T_{ij}$ is the probability that a customer chooses to travel by the vehicle $i$. The system is then in equilibrium if and only if

$$
m_u := \sum_{ij} T_{ij} \lambda_{ij}(u_{ij}) > 0.
$$
which combined with equation \[9\] and the fact that \(c = 0\) implies that
\[
0 = -\lambda_i + \sum_j \lambda_{ij} H(v_j). \tag{11}
\]
To arrive at \[11\], we used the fact that in the stationary equilibrium, \(v_i\) and \(c_i\) are constants and hence, there is no dependence on \(t - T_{ij}\).

Recall that for every equilibrium solution, we require \(v > 0\) and thus \(H(v_i) = 1, \forall i \in \mathcal{N}\). Therefore, a necessary condition for the existence of equilibria is that the prices \(u\) are chosen such that
\[
0 = -\lambda_i + \sum_j \lambda_{ij}, \quad \forall i \in \mathcal{N},
\]
which proves the first statement.

We now want to verify that the fleet size is large enough to maintain an equilibrium flow. Recall the fleet size dynamics \(\dot{n}(t)\) when \(c = 0\) and \(v > 0\) in \[6\]. Observe that to satisfy \(v > 0\), one needs to have a fleet size of at least \(\sum_{ij} T_{ij} \lambda_{ij}(u_{ij})\) vehicles which is the definition of \(m_u\). This, mixed with its invariant property \((\dot{n} = 0)\), proves the claim. Conversely, if \(m < m_u\) no equilibrium exists.

**Lemma 1 (Existence of an equilibrium).** The set \(\mathcal{U}\) is never empty, hence, at least one equilibrium exists.

**Proof:** We use the fact that there exists a price \(u_{ij}^{\text{max}}\) for which \(\lambda_{ij}(u_{ij}^{\text{max}}) = 0\) for all \(i, j \in \mathcal{N}\). Then, setting \(u = u^{\text{max}}\), implies that an equilibrium exists. This strategy means that we are not providing service to any request, nevertheless the equilibrium exists as we are unable to have an invariant fleet size.

**Lemma 2 (Infinite number of equilibria).** If there is a positive demand tour in the graph, then there exists an infinite number of price vectors \(u\) which can steer the system to an equilibrium point.

**Proof:** Assume that there exists at least one Eulerian tour (or cycle) in the graph for which \(\lambda_i^0 > 0\) for all \((i, j) \in \text{cycle}\). Then, let \(\lambda_{\text{cycle}} = \{\lambda_{ij}^0 | (i, j) \in \text{cycle}\}\) and the minimum rate on that tour be \(\lambda_{\text{min}}^0 = \min_{i \in \text{cycle}} \lambda_{ij}^0\). Then by setting \(u_{ij} = 0\) for all \((i, j) \not\in \text{cycle}\), we can express the equilibrium condition as
\[
\sum_{j:(i, j) \in \text{cycle}} \lambda_{ij}(u_{ij}) - \lambda_{ji}(u_{ij}) = 0, \quad \forall i : (i, j) \in \text{cycle}. \tag{12}
\]
Now, we use the fact that \(\lambda_{ij}(u_{ij})\) is a monotonically decreasing function and we focus on \((i, j) \in \text{cycle}\).

Hence for all \(\lambda_{ij}(u_{ij}) > \lambda_{\text{cycle}}^0\) we can find a \(u_{ij}\) such that \(\lambda_{ij}(u_{ij}) = \lambda_{\text{cycle}}^0\). Then, extending this for higher prices on \(\lambda_{\text{cycle}}^0\) and using the same argument as before, we show that there exists a pricing strategy \(u\) for which we can obtain an equilibrium with a tour demand rate with any value in the range \((0, \lambda_{\text{cycle}}^{\text{min}})\).

**Corollary 1 (Minimum number of vehicles in equilibria).** The minimum number of vehicles to operate in an equilibrium induced by policy \(u\) is at least
\[
m > m := \min_u m_u
\]
where \(m_u := \sum_{ij} T_{ij} \lambda_{ij}(u_{ij})\).

**Proof:** This result follows directly from the last argument in the proof of Theorem 1.

**B. Stability**

In this section we study local stability of the equilibria presented in the previous subsection. As an example, we look at cases when a disruptive change happens to the system, either because of an increase in customers or a decrease in the availability of vehicles. Let \(u \in \mathcal{U}\) and assume \(m_u > m\). Then, we define the set of equilibria as
\[
\Upsilon_u := \{(c, \nu) \in \mathbb{R}^{2N} | c_i = 0, v_i > 0, \forall i \in \mathcal{N},
\] and \(\sum_i v_i = M - m_u\). \tag{13}
\]
**Definition 1 (Locally asymptotically stable).** We say that a set of equilibria \(\Upsilon_u\) is locally asymptotically stable if for any equilibrium \((c, \nu) \in \Upsilon_u\), there exists a neighborhood \(B^U_u(c, \nu)\) such that every evolution of the model \[5\] starting at \((c(\tau), \nu(\tau)) = (c, \nu)\) and with \((c(0), \nu(0)) \in B^U_u(c, \nu)\) has a limit which belongs to the equilibrium set \(\Upsilon_u\) i.e., \((c(\tau) \rightarrow +\infty c(t), \nu(\tau) \rightarrow +\infty v(t)) \in \Upsilon_u\) where \(\tau \in [-m_{\text{cycle}}, T_{ij}, 0)\) and
\[
B^U_u(c, \nu) := \{(c, \nu) \in \mathbb{R}^{2N} | c_i > 0, v_i = v_i, \forall i \in \mathcal{N},
\] and \(||(c - c^0)|| < \delta\}. \tag{14}
\]
**Theorem 2 (Stability of the equilibria).** Let \(u \in \mathcal{U}\) and assume \(m_u > m\); then, the set of equilibria \(\Upsilon_u\) is locally asymptotically stable.

**Proof:** Provided in the Appendix.

**IV. OPTIMAL STRATEGIES**

In this section, we present an optimization framework to find optimal prices given endogenous demand rates. This model aims to maximize the revenue of an AMoD provider while ensuring load balancing of clients and vehicles. We then turn to a model which uses a rebalancing formulation to ensure load balancing, without the need of price adjustments. Finally, we combine these two formulations into a joint model.

**A. Optimal Pricing**

We are interested in finding the best pricing policy that ensures the existence of an equilibrium \([7]\). Hence, we define the feasible set of the pricing problem to be
\[
\mathcal{F} = \left\{ u : \sum_i (\lambda_{ij}(u_{ij}) - \lambda_{ji}(u_{ji})) = 0, \quad \forall j \in \mathcal{N}, \quad u \in [1, u^{\text{max}}]\right\}.
\]
Then, we can define the profit maximization problem as
\[
\max_{u \in \mathcal{F}} \sum_{ij} \lambda_{ij}(u_{ij}) u_{ij} p_{ij} - c_i^0 \lambda_{ij}(u_{ij}) - c_i^0 (\lambda_{ij}^0(u_{ij}) - \lambda_{ij}(u_{ij})), \tag{15}
\]
where $\lambda_{ij}(u_{ij})u_{ij}p_{ij}^0$ and $c_{ij}^o$ are the total revenue and the operational cost of request $i$ to $j$, respectively; and $c^r$ is the penalty that the AMoD service incurs when a customer exists the platform because of a high price.

Note that if the functions $J_{ij}(u_{ij}) := \lambda_{ij}(u_{ij})u_{ij}p_{ij}^0 - c_{ij}^o \lambda_{ij}(u_{ij}) - c^r(\lambda_{ij}(u_{ij}) - \lambda_{ij}(u_{ij}))$ are concave in the range of $[1, u_{ij}^{\max}]$, then the optimization problem is tractable (we maximize over a concave function with linear equality constraints). To ensure the concavity of the cost function $J_{ij}$ we need its second derivative to satisfy

$$j_{ij} \leq 0 \implies \lambda_{ij}(u_{ij}) \leq -\frac{2}{u_{ij}p_{ij}^0 - c_{ij}^o} \lambda_{ij}(u_{ij}).$$

(16)

Recall that by Assumption 1 ($\lambda_{ij}$ is monotonically decreasing), $\lambda_{ij} < 0$. Hence, for any linear demand function, the problem becomes tractable.

B. Optimal Rebalancing

We use the planning rebalancing model developed in [13]. In this setting, we aim to find a static rebalancing policy that reaches an equilibrium. Let the rebalancing flow be $r_{ij}$, that is, the rate at which empty vehicles flow from $i$ to $j$. To solve the problem we use the following Linear Program (LP) that minimizes the empty travel time and seeks to equate the inflow and outflow of vehicles at each region by using $N^2$ variables

$$\min_{r \geq 0} \sum_{ij} T_{ij} r_{ij}$$

(17a)

subject to

$$\sum_{i} \lambda_{ij}^0 + r_{ij} - \lambda_{ji}^0 - r_{ji} = 0, \forall j \in \mathbb{N}.$$  

(17b)

Notice that in this case we use $\lambda_{ij}^0$ instead of $\lambda_{ij}(u_{ij})$ as we do not consider the possibility of decreasing the demand by using price incentives. This LP is always feasible as one can always choose $r_{ij} = \lambda_{ij}^0 > 0$ for all $i, j \in \mathbb{N}$ which satisfies the set of constraints (17b). All the results presented in Section III hold for this problem as well and are studied in [13].

C. Joint Pricing and Rebalancing

We are interested in choosing the best policy which leverages different decisions that the mobility-on-demands providers face. In particular, we will like to optimize the pricing, re-balancing and sizing problem. Then, we can write the planning optimization problem as,

$$\max_{\lambda, r, u, m} \sum_{ij} \lambda_{ij}(u_{ij})u_{ij}p_{ij}^0 - c_{ij}^o \lambda_{ij}(u_{ij}) - c^r(\lambda_{ij}(u_{ij}) - \lambda_{ij}(u_{ij})) - c^r(r_{ij}T_{ij}) - c^r m$$

s.t.

$$\sum_{i} \lambda_{ij}(u_{ij}) + r_{ij} - \lambda_{ji}(u_{ji}) - r_{ji} = 0, \forall j \in \mathbb{N};$$

$$\sum_{ij} T_{ij}(\lambda_{ij}(u_{ij}) + r_{ij}) \leq m, \forall j \in \mathbb{N};$$

$$u \in [1, u_{ij}^{\max}],$$

(18b)

(18c)

(18d)

where $c^r$ and $c^f$ are the cost of rebalancing and the cost of owning and maintaining a vehicle per unit time, respectively. Note that to ensure that solving (18) reaches a global maximum, we must validate that (16) holds for $u \in [1, u_{ij}^{\max}]$.

Note that this problem, if solvable, yields a solution with higher profits than the individual formulations of pricing (15) and (17), or the sequential approach of solving first the rebalancing problem and then adjusting prices. This happens given that the problem is jointly solving for $u$ and $r$ considering simultaneously the full objective of the profit maximization problem (18a).

V. Experiments

We carry out two case studies to assess the benefits of solving the joint problem of pricing and rebalancing over other approaches. Our first experiment uses a fictitious transportation network to analyze sensitivities with respect to respect to the network size. The second one consists of a data-driven case study using historical data from New York City. We report empirical results of the achievable profit of the AMoD system when solving the problem using the different methodologies presented in Table I.

**TABLE I: Different policies evaluated to plan the operation of an AMoD system.**

| Policy | Type | Formulation |
|--------|------|-------------|
| $P_{ij}$ | Individual | [15] |
| $R_{ij}$ | Individual | [16] |
| $R_{ij} \rightarrow P_{ij}$ | Sequential | [17] then [15] |
| $P_{ij} + R_{ij}$ | Joint with fixed price by origin | [18] with $u_{ij} = u_{ij}^k$ |
| $P_{ij} + R_{ij}$ | Joint | [18] |

We begin with the individual the policies $P_{ij}$ and $R_{ij}$ to see the equilibrium under a policy or rebalancing strategy. We then turn to a sequential approach $R_{ij} \rightarrow P_{ij}$ to solve the problem. Our motivation for this methodology comes from the fact that many corporations tend to have separate pricing and rebalancing departments, which would result in solving the joint problem sequentially. Note that the sequential policy $P_{ij} \rightarrow R_{ij}$ is not included because once the pricing problem is solved, the system is at equilibrium and the rebalancing problem becomes trivial (i.e., $r = 0$).

Finally, the joint with fixed prices by origin policy $P_{ij} + R_{ij}$ is motivated by the fact that current MoD services only use the origin (not the destination) when setting surge prices (price multipliers) [16], [17].

Note that in order to have a tractable solution for formulations (15) and (17), we require a function satisfying (16). To achieve this, we assume a linear demand (willingness-to-pay) function, specifically we let

$$\lambda_{ij}(u_{ij}) = \frac{\lambda_{ij}^0}{u_{ij}^{\max}}(u_{ij}^{\max} - u_{ij}),$$

(19)

where we set $u_{ij}^{\max} = 4$ as suggested in [17]. Hence, by using this linear demand function, we get a tractable Quadratic Program (QP) with linear constraints. Arguably, linear demand functions may not be as accurate as desired for realistic implementations of this model. However, using linear functions allows us to recover a global maximum solution to the problem and assess the potential benefits that joint policies may achieve compared to other strategies.

For both experiments we let the base price be proportional to the travel time using $p_{ij}^0 = 0.5T_{ij}$, where $0.5$ is the average price a user pays in dollars per minute of taxi.
ride reported in [18]. Additionally, we let the operation and rebalancing cost per kilometer be $c^o = c^r = $0.72 as suggested in [19]; the lost customer cost $c^e$ is equal to $5$, and we set the regularal parameter on the fleet size to be $c^f = 1 \times 10^{-10}$.

A. Uniform Demand

We compare the solution of the different methods for a network with random uniform demands. For each strategy, we let the base demand be $\lambda_{ij}^0 \sim U(0, 4)$ and travel time between regions be $T_{ij} \sim U(0, 40)$. Then, we solve the problem for networks with a number of regions ranging between 0 and 60.

Figure 3a shows the value of the cost function (18a) for each methodology. Moreover, in Figure 3b we observe the relative deviation in profits for the solution of each strategy against the joint pricing and rebalancing solution. We see that as the number of regions increases, the deviation converges to a stable value. To explain this phenomenon, we define the potential of region $i$ to be the load balance deviation when no pricing or rebalancing policy is applied, namely, $\zeta_i = \sum_j \lambda_{ij}^0 - \lambda_{ij}^0$. Then, since we draw samples from the same uniform distribution to assign all $\lambda_{ij} \ \forall i,j \in \mathcal{N}$, the expected value of $\zeta_i$ is equal to zero for all $i$. Hence, this convergence behavior is simply a direct implication of the law of large numbers. Note that, for the same reason, the individual policy $P_{ij}$ converges to zero.

B. New York City Case Study

We perform a case study of New York city using the data available at [20]. Specifically, we analyze the data set of High Volume For-Hire Vehicle Trip Records of November 2019 [20]. In order to analyze stable distributions of trips in the network, we filter the data to consider only working days (Monday to Friday). Then, we focus on four time slots: Morning Peak (AM) from 7:00-10:00 hrs, Noon (MD) from 12:00-15:00 hrs, Afternoon Peak (PM) from 17:00-20:00 hrs and Night (NT) from 00:00-3:00 hrs. For every time window in November 2019, we collect data on origin-destination pairs and travel times of every trip. Then, we compute the average hourly demand and travel times, and we use these values to preform our analysis and test the different solutions.

Table II shows the deviation in profits (in percentage terms) between the different approaches and the joint formulation. As a reminder, use Table I summarizes all policy definitions. We observe that the joint method outperforms all the other methods in the range from 5% to 40%, highlighting the benefit of solving this problem using a joint strategy. In particular, we observe that each of the individual strategies performs on average worse than strategies that optimize both pricing and rebalancing. Also, it is relevant to stress the 5% deviation of the policy with fixed surge price by origin, as it matches our expectations of the relevance of considering the destination when pricing.

To better understand the different approaches, we generated plots of the pricing distribution and trend. Figure 4 shows histograms comparing the value of the solution $u^*$ for the individual pricing policy and the joint strategy. As expected, we observe the distribution of the individual approach to have higher variance than the joint. This happens given the hard constraint to reach an equilibrium. When no rebalancing is considered as in $P_{ij}$ the policy chooses prices to ensure $u \in \mathcal{F}$. In contrast, when solving the joint problem, the solution leverages rebalancing and pricing and gives the pricing decision more flexibility to concentrate to pick values that maximize profits.

Figure 5 plots prices against $\zeta_i$ (the potential of region $i$). Recall that parameter $\zeta_i$ indicates excess demand for positive values (i.e., more costumers than vehicles) and excess supply for negative values (i.e., more vehicles than costumers). For the problem with unique surge prices we plot values by origin. For the joint problem we plot the demand-weighted average price by origins. Just as we would expect, there is a positive trend between these variables. The algorithm lowers prices when there is excess supply to incentivize users to request rides, and increases the price when there is excess demand. Note that the pattern is stronger for busier times (AM and PM). Finally, we quantify how relevant is the pricing relative to the rebalancing component when balancing the load of the system. Letting $f^*$ and $u^*$ be the solution of (18), we define a load dispersion metric as follows $\zeta_0 = \frac{1}{n} \sum_i |\sum_j \lambda_{ij}^0 - \lambda_{ij}^0|$ when nothing is applied,
framework to operate AMoD systems. Finally, we would like to solve the problem from a welfare maximization perspective rather than from the profit maximization and compare the results.

REFERENCES

[1] S. Wollenstein-Betech, A. Houshmand, M. Salazar, M. Pavone, C. G. Cassandras, and I. C. Paschalidis, “Congestion-aware routing and rebalancing of autonomous mobility-on-demand systems in mixed traffic,” 2020, Submitted to IEEE ITSC Conference 2020. [Online]. Available: https://arxiv.org/pdf/2003.04335.pdf

[2] S. Banerjee, K. Johan, and C. Riqueur, “Pricing in ride-sharing platforms: A queueing-theoretic approach,” in Proceedings of the Sixteenth ACM Conference on Economics and Computation, 2015, pp. 49–69.

[3] B. Turan, R. Pedarsani, and M. Alizadeh, “Dynamic pricing and management for electric autonomous mobility on demand systems using reinforcement learning,” 2019.

[4] K. Bimpikis, O. Candogan, and D. Saban, “Spatial pricing in ride-sharing networks,” Operations Research, vol. 67, no. 3, pp. 744–769, 2019.

[5] R. M. Swaszek and C. G. Cassandras, “Load balancing in mobility-on-demand systems: Reallocation via parametric control using concurrent estimation,” in 2019 IEEE Intelligent Transportation Systems Conference (ITSC), IEEE, 2019, pp. 2148–2153.

[6] S. Hörl, C. Ruch, F. Becker, E. Frazzoli, and K. W. Axhausen, “Fleet control algorithms for automated mobility: A simulation assessment for Zurich,” 2018.

[7] M. W. Levin, K. M. Kockelman, S. D. Boyles, and T. Li, “A general framework for modeling shared autonomous vehicles with dynamic network-loading and dynamic ride-sharing application,” Computers, Environment and Urban Systems, vol. 64, pp. 373 – 383, 2017.

[8] R. Zhang and M. Pavone, “Control of robotic Mobility-on-Demand systems: A queueing-theoretical perspective,” vol. 35, no. 1–3, pp. 186–203, 2016.

[9] R. Iglesias, F. Rossi, R. Zhang, and M. Pavone, “A BCMP network approach to modeling and controlling Autonomous Mobility-on-Demand systems,” 2016.

[10] M. Pavone, S. L. Smith, E. Frazzoli, and D. Rus, “Robotic load balancing for Mobility-on-Demand systems,” vol. 31, no. 7, pp. 839–854, 2012.

[11] F. Rossi, R. Zhang, Y. Hindy, and M. Pavone, “Routing autonomous vehicles in congested transportation networks: Structural properties and coordination algorithms,” vol. 42, no. 7, pp. 1427–1442, 2018.

[12] M. Salazar, M. Tsao, I. Aguiair, M. Schiffer, and M. Pavone, “A congestion-aware routing scheme for autonomous mobility-on-demand systems,” 2019.

[13] M. Pavone, S. L. Smith, E. Frazzoli, and D. Rus, “Robotic load balancing for mobility-on-demand systems,” The International Journal of Robotics Research, vol. 31, no. 7, pp. 839–854, 2012.

[14] A. F. Filippov, Differential equations with discontinuous righthand sides: control systems. Springer Science & Business Media, 2013, vol. 18.

[15] G. Haddad, “Monotone trajectories of differential inclusions and functional differential inclusions with memory,” Israel journal of mathematics, vol. 39, no. 1-2, pp. 83–100, 1981.

[16] L. Chen, A. Mislove, and C. Wilson, “Peeking beneath the hood of uber,” in Proceedings of the 2015 internet measurement conference, 2015, pp. 495–508.

[17] P. Cohen, R. Hahn, J. Hall, S. Levitt, and R. Metcalfe, “Using big data to estimate consumer surplus: The case of uber,” National Bureau of Economic Research, Tech. Rep., 2016.

[18] “Taxi fares: How much does a ride cost?” [Online]. Available: https://www.seattle.gov/your-rights-as-a-customer/file-a-complaint/taxi-for-hire-and-tnc-complaints/taxi-fares-how-much-does-a-ride-cost

[19] P. M. Bsch, F. Becker, H. Becker, and K. W. Axhausen, “Cost-based analysis of autonomous mobility services,” Transport Policy, vol. 64, pp. 76 – 91, 2018.

[20] “Tlc trip record data.” [Online]. Available: https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page

[21] G. Sun, C. G. Cassandras, Y. Wang, C. G. Papayioutou, and G. F. Riley, “Perturbation analysis and optimization of stochastic flow networks,” IEEE Transactions on Automatic Control, vol. 49, no. 12, pp. 2143–2159, 2004.

[22] C. G. Cassandras and S. Lafortune, Introduction to Discrete Event Systems, 2nd ed. Springer, 2010.
Proof of Theorem 2

We start by showing that \( v_i(\tau) > c_i(\tau) \) for \( \tau \in [-\max_{i,j} T_{ij}, t] \), which will serve as a key element to the analysis. To do so, we first assume \( v_i(\tau) > 0 \) for all \( i \in \mathcal{N} \), and observe the system dynamics (5) at time \( t \) are

\[
\dot{c}_i(t) = (\lambda_i - \mu_i) H(c_i),  
\]

(20a)

\[
\dot{v}_i(t) = -\lambda_i + (\lambda_i - \mu_i) H(c_i) + \sum_j (\lambda_{ji} - \mu_j) H(c^j),  
\]

(20b)

\[
= (\lambda_i - \mu_i) H(c_i) - \sum_j (\lambda_{ji} - \mu_j) H(c^j),  
\]

(20c)

\[
\geq (\lambda_i - \mu_i) H(c_i),  
\]

(20d)

\[
= \dot{c}_i(t),  
\]

(20e)

where all the \( H(v_i) \) in (5) are replaced with 1 since we assume that \( v_i(\tau) > 0 \). The step (20a) is due to the fact that the system is at equilibrium, i.e., \( \sum_j \lambda_{ji} - \lambda_i = 0 \), \( \forall i \in \mathcal{N} \), and the step (20d) is a result of \( \mu_i > \lambda_i \) which means that \( \lambda_i - \mu_i < 0 \). Given that \( \dot{v}_i(t) \geq \dot{c}_i(t) \) and the fact that at the starting point \( v > c \) (i.e., \( v > 0 \)), we conclude that \( v_i(\tau) > c_i(\tau) \) for \( \tau \in [-\max_{i,j} T_{ij}, t] \) and \( i \in \mathcal{N} \).

Two important consequences of this result are that \( c_i \) always reaches 0 before its corresponding \( v_i \), and the vehicle time derivative \( \dot{v}_i \) is greater than or equal to 0 after \( c_i \) has reached 0. This follows by observing that all the terms in (5b) are all positive when \( c_i = 0 \).

Now we are in a position to show that \( v(t) > 0 \) for \( t \geq 0 \). We do this by combining the second consequence in the previous paragraph with the assumption that the initial state of the system is \((0, \bar{v})\) and the fact that \( \dot{v}_i(t) > \dot{c}_i(t) \). Thus, since \( v(t) > 0 \) for \( t \geq 0 \), we have that \( \dot{c}_i(t) = (\lambda_i - \mu_i) H(c_i) \leq 0 \) which will implies that \( c_i \) will for sure be 0 when \( t \geq T_{\max} \) where \( T_{\max} := \max_{i,j} \{ c_i(0) / (\mu_i - \lambda_i) \} \) \( \forall i \in \mathcal{N} \) and which implies that \( \lim_{t \to +\infty} c_i(t) = 0 \) since both \( \dot{c}_i \) and \( c_i \) will be equal to 0 for all \( i \in \mathcal{N} \).

To show that \( \lim_{t \to +\infty} v(t) = \bar{v} \) we use the fact that \( c_i = \dot{c}_i = 0 \) for \( t > 0 \) and insert this into the vehicle dynamics in (5b) obtaining \( \dot{v}_i(t) = \lambda_i v_i + \sum_j (\lambda_{ji} H(v^j) - (\lambda_{ji} - \mu_j) H(c^j)). \) Since, \( c_i(t) = 0 \) we observe that after \( T_{\max} + \max_{i,j} T_{ij} \) time units \( H(c^j) \) will be equal to zero and therefore \( \dot{v}(t) = 0 \) for \( t > T_{\max} + \max_{i,j} T_{ij} \). Moreover, since \( \dot{v}_i(t) = 0 \) the \( \lim_{t \to +\infty} v_i(t) \) exists and can be retrieved using \( v_i(t) = v_i(0) + \int_0^t \dot{v}_i(s) ds = v_i(0) + c_i(0). \) Given that we show that \( v_i(0) > c_i(0) \), we conclude that \( \lim_{t \to +\infty} v_i(t) > 0 \). The property \( \lim_{t \to +\infty} v_i(t) > 0 \) holds straight from Proposition 1

Finally, to characterize the ball \( B_{\delta}^U \) we set \( \psi := v_1 \sin(\pi/4) \) and \( \psi_{\min} := \min_i \psi_i \) (see Fig. 6). Then, for \( \delta = \psi_{\min} \) any path of the system (5) starting at \( (c(\tau), v(\tau)) \) for \( \tau \in [-\max_{i,j} T_{ij}, 0] \) and satisfying \( (c(0), v(0)) \in B_{\delta}^U \) has a limit which belongs to the equilibrium set \( \mathcal{E} \).

Fig. 6: Sketch of a variable of the initial solution \((c, \bar{v})\) along with its neighborhood \( B_{\delta}^U \). Shaded in grey is the feasible region \( (c_i < v_i) \).