

# Load Balancing for Mobility-on-Demand Systems

Marco Pavone  
Jet Propulsion Laboratory, Caltech  
Pasadena, CA 91109 USA  
marco.pavone@jpl.nasa.gov

Stephen L. Smith  
University of Waterloo  
Waterloo ON, N2L 3G1 Canada  
stephen.smith@uwaterloo.ca

Emilio Frazzoli Daniela Rus  
Massachusetts Institute of Technology  
Cambridge, MA 02139 USA  
frazzoli@mit.edu, rus@csail.mit.edu

**Abstract**—In this paper we develop methods for maximizing the throughput of a mobility-on-demand urban transportation system. We consider a finite group of shared vehicles, located at a set of stations. Users arrive at the stations, pick-up vehicles, and drive (or are driven) to their destination station where they drop-off the vehicle. When some origins and destinations are more popular than others, the system will inevitably become out of balance: Vehicles will build up at some stations, and become depleted at others. We propose a robotic solution to this rebalancing problem that involves empty robotic vehicles autonomously driving between stations.

We develop a rebalancing policy that minimizes the number of vehicles performing rebalancing trips. To do this, we utilize a fluid model for the customers and vehicles in the system. The model takes the form of a set of nonlinear time-delay differential equations. We then show that the optimal rebalancing policy can be found as the solution to a linear program. By analyzing the dynamical system model, we show that every station reaches an equilibrium in which there are excess vehicles and no waiting customers. We use this solution to develop a real-time rebalancing policy which can operate in highly variable environments. We verify policy performance in a simulated mobility-on-demand environment with stochastic features found in real-world urban transportation networks.

## I. INTRODUCTION

In the past century, private automobiles have dramatically changed the concept of *personal urban mobility*, by enabling fast and anytime point-to-point travel within large cities. In 2001, personal urban mobility in the US resulted in more than 3.5 trillion urban miles traveled by private cars, representing 75% of total car travel in the US [16]. This figure, coupled with the fact that by 2030 the total population living in urban areas will jump from the current 40% to more than 60% [15], implies that the “demand for personal urban mobility” will increase to formidable levels. The demand for roads and parking spaces will dramatically increase, while the available urban land will continue to decrease. The result is that private automobiles are an *unsustainable* solution for the future of personal mobility in dense urban environments. To cope with this problem, a paradigm shift is emerging to replace the outdated policy of infrastructure augmentation with personal urban mobility leveraging robotics and automation. The challenge is to ensure the same benefits of privately-owned cars without requiring additional roads and parking spaces.

One of the leading emerging paradigms for future urban mobility systems is one-way vehicle sharing, which effectively merges private and public mobility, and directly targets the problems of parking spaces and current low vehicle utilization rates. Arguably, the most promising approach within

this paradigm is represented by *Mobility-on-Demand* (MOD) systems [14], which provide stacks and racks of light electric vehicles at closely spaced intervals throughout a city: when a person wants to go somewhere, he/she simply walks to the nearest rack, swipes a card to pick up a vehicle, drives it to the rack nearest to his/her destination, and drops it off (see Figure 1 for an illustration of mobility-on-demand). Large-scale systems employing traditional, non-electric bicycles have already demonstrated the feasibility of mobility-on-demand in several cities throughout Europe, e.g., Paris, Lyon, Milano, Trento, Zurich [13].

However, sharing has its drawbacks. When some origins and some destinations are more popular than others, the system will inevitably become out of balance: Vehicles will build up at some stations, and become depleted at others. In this paper we propose a *robotic solution for vehicle rebalancing* in MOD systems, whereby the shared vehicles autonomously drive from a delivery location to the next pick-up location. Rebalancing through autonomously driving vehicles has the clear potential of eliminating imbalances within the transportation network, and effectively adds an additional dimension to MOD systems by introducing autonomy in the design space. In the recent past, considerable efforts have been devoted to the problem of autonomous driving, and substantial progress has been made (see, for example, [4, 3]). However, there are virtually no tools to address the system-level problems arising at the interface between robotics and transportation science: How should one pre-position vehicles in order to anticipate future demand? Is it possible to characterize optimal, real-time rebalancing policies? How many vehicles are needed to achieve a certain quality of service (e.g., a desired average waiting time for the customers)? The purpose of this paper is to develop an approach that provides rigorous answers to such questions.

Even though rebalancing in MOD systems is an entirely new problem within the realm of transportation networks, it has many characteristics in common with the well-known Dynamic Traffic Assignment (DTA) problem [12, 8, 20, 19]. In this problem, one seeks to “optimize” the time varying flows on each arc of a transportation network, taking into account congestion effects along arcs and at nodes [8]. DTA models mainly differ in the methods used to capture the time-varying nature of supply and demand processes, and can be broadly divided into four categories [19]: (i) mathematical programming (e.g., [12, 20]), (ii) optimal control (e.g., [8]), (iii) variational inequality (e.g., [9]), and (iv) simulation-based (e.g., [1]). The key difference between MOD systems

and the DTA problem is that we are optimizing over the empty vehicle trips (i.e., the rebalancing trips) rather than the passenger carrying trips. Rebalancing in MOD systems is also related to Dynamic one-to-one Pick-up and Delivery problems (DPDPs), where dynamically-generated passengers must be transported from a pick-up site to a delivery site by a fleet of vehicles. DPDPs can be divided into three main categories [2]: (i) Dynamic Stacker Crane Problem, where the vehicles have unit capacity, (ii) Dynamic Vehicle Routing Problem with Pickups and Deliveries, where the vehicles can transport more than one request, and (iii) Dynamic Dial-a-Ride Problem, where additional constraints such as time windows are considered. Excellent surveys on heuristics, metaheuristics and online algorithms for DPDPs can be found in [2] and [17], while analysis specifically tailored to the structural properties of transportation-on-demand systems can be found in [18]. The key difference from DPDP problems is that in the MOD system there are a finite number of pick-up and delivery sites, but the vehicles are not aware of the destination of newly arrived customers. Finally, our problem is also related to dynamic load balancing in distributed computing systems [6, 5]. However, these problems are less constrained since the demands (i.e., jobs) do not need to wait for a “vehicle” to move across multiple processors.

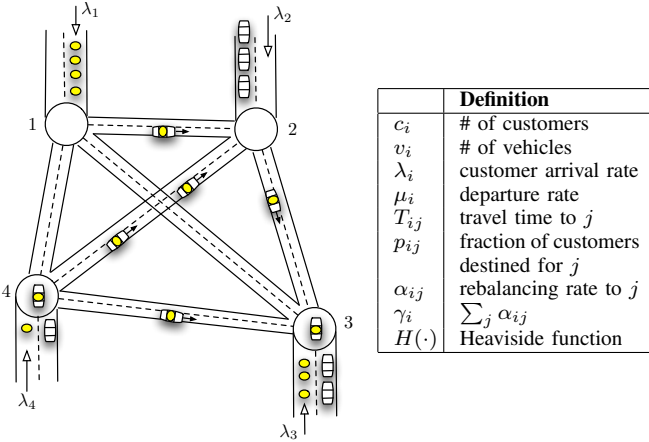


Fig. 1. At each station there is a queue of customers (yellow dots) and a queue of vehicles (represented by small car icons). The customer at the head of queue enters the vehicle at the head of the queue. This is shown in the circles at stations 3 and 4. Notice that at station 1 there are no vehicles, and at station 2 there are no customers. In rebalancing, we send empty vehicles from station 2 to station 1.

This paper is structured as follows. In Section II we present a fluid model for MOD systems, and we formally state the rebalancing problem. In Section III we (i) study well-posedness and equilibria of the fluid model, and we show that without rebalancing vehicles the system is unstable (i.e., at some stations the number of waiting customers will grow without bound); (ii) determine the minimum number of vehicles needed to meet the customer demand; and, (iii) show how to optimally route the rebalancing vehicles across the transportation network. In Section IV we tackle the problem of *robust* rebalancing, and we study policies to handle, for

example, a sudden “burst” of customers arriving at the stations. The results in Sections III and IV lead to a robust, real-time policy for vehicle rebalancing, which is presented in Section V and which is thoroughly evaluated through simulation experiments in Section VI.

## II. MODELING THE MOBILITY-ON-DEMAND SYSTEM

In this paper, we use a *fluidic approach* to model Mobility-on-Demand systems.<sup>1</sup> Consider  $n$  stations defined over an extended geographical area (see Figure 1). We denote the set of stations with  $\mathcal{N}$ . In this model, the number of customers and vehicles are represented by real numbers. Customers arrive at station  $i$  at a constant rate  $\lambda_i \in \mathbb{R}_{>0}$ . The number of customers at station  $i$  at time  $t$  is  $c_i(t) \in \mathbb{R}_{\geq 0}$ , and the number of vehicles waiting idle at station  $i$  at time  $t$  is  $v_i(t) \in \mathbb{R}_{\geq 0}$ . The total number of vehicles in the system is  $V \in \mathbb{R}_{>0}$ . The fraction of customers at station  $i$  whose destination is station  $j$  is  $p_{ij}$  (where  $p_{ij} \in \mathbb{R}_{\geq 0}$ ,  $p_{ii} = 0$ , and  $\sum_j p_{ij} = 1$ ). The travel time from station  $i$  to station  $j$  as  $T_{ij} \in \mathbb{R}_{\geq 0}$ . When there are both customers and vehicles at station  $i$  (i.e.,  $c_i(t) > 0$  and  $v_i(t) > 0$ ), then the rate at which customers (and hence vehicles) leave station  $i$  is  $\mu_i$ ; when, instead,  $c_i(t) = 0$  but  $v_i(t) > 0$  the departure rate is 0. A necessary condition for the total number of customers at station  $i$  to remain bounded is that  $\mu_i \geq \lambda_i$ ; we will assume  $\mu_i \geq \lambda_i$  throughout the paper.

In order to rebalance the number of vehicles at each station, empty vehicles will be sent between stations. The rate at which station  $i$  sends empty vehicles to station  $j$  is denoted by  $\alpha_{ij} \in \mathbb{R}_{\geq 0}$  and the total rate at which station  $i$  sends empty vehicles is  $\gamma_i := \sum_j \alpha_{ij}$ , where  $\alpha_{ii} = 0$ . We let  $\alpha$  denote the matrix with entries given by  $\alpha_{ij}$ . The notation is summarized in Figure 1.

We are now in a position to write the differential equations governing the evolution of the number of vehicles and customers at each station. In order to write the expressions more compactly, we introduce the following shorthand notation:

$$v_i := v_i(t) \quad c_i := c_i(t) \quad v_j^i := v_j(t - T_{ji}) \quad c_j^i := c_j(t - T_{ji}).$$

Then, we can write the customer dynamics at station  $i$  as

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

Defining the Heaviside function as

$$H(x) := \begin{cases} 1, & \text{if } x > 0, \\ 0, & \text{otherwise,} \end{cases}$$

the customer dynamics can be written as

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

<sup>1</sup>Our fluid model is intended to serve as an approximation of a corresponding stochastic queueing model, where customers enter the system according to a Poisson process. This can be formalized by showing that the fluid model arises as the limit of a sequence of appropriately scaled queueing models. Such analysis will be presented in a forthcoming paper. In this paper, the relation between the fluidic approximation and the queueing model will be discussed through simulation in Section VI.

The rate of change of vehicles at station  $i$  can be written as the sum of four components:

- (i) the rate at which customer-carrying vehicles depart station  $i$ :

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

which can be written more compactly as  $-\lambda_i H(v_i) + (\lambda_i - \mu_i) H(c_i) H(v_i)$ .

- (ii) the rate at which customer-carrying vehicles arrive at station  $i$ :

$$\sum_{j \neq i} p_{ji} \left( \lambda_j H(v_j^i) + (\lambda_j - \mu_j) H(c_j^i) H(v_j^i) \right)$$

- (iii) the rate at which empty (re-balancing) vehicles depart station  $i$ , given by  $\gamma_i H(v_i)$ .

- (iv) the rate at which empty (re-balancing) vehicles arrive at station  $i$ , given by  $\sum_{j \neq i} \alpha_{ji} H(v_j^i)$ .

Putting everything together, we can write a set of nonlinear, time-delay, differential equations describing the evolution of customers and vehicles in the system as

$$\begin{aligned} \dot{c}_i &= \lambda_i (1 - H(v_i)) + (\lambda_i - \mu_i) H(c_i) H(v_i), \\ \dot{v}_i &= -\lambda_i H(v_i) + (\lambda_i - \mu_i) H(c_i) H(v_i) + \sum_{j \neq i} p_{ji} \left( \lambda_j H(v_j^i) \right. \\ &\quad \left. - (\lambda_j - \mu_j) H(c_j^i) H(v_j^i) \right) - \gamma_i H(v_i) + \sum_{j \neq i} \alpha_{ji} H(v_j^i). \end{aligned} \quad (1)$$

The initial conditions satisfy  $c_i(\tau) = 0$ ,  $v_i(\tau) = 0$  for  $\tau \in [-\max_{i,j} T_{ij}, 0)$ , and  $c_i(0) \in \mathbb{R}_{\geq 0}$ ,  $v_i(0) \in \mathbb{R}_{\geq 0}$  (with  $v_i(0) > 0$  for at least one  $i \in \mathcal{N}$ ). Accordingly, we write  $V = \sum_i v_i(0)$  (we will show in Proposition 3.1 that the number of vehicles is invariant along system trajectories).

The problem we wish to solve is as follows: find a rebalancing assignment  $\alpha$  that minimizes the number of empty vehicles traveling in the network and ensures the existence of stable equilibria for model (1).

### III. WELL-POSEDNESS, EQUILIBRIA, AND OPTIMAL REBALANCING

In this section we first discuss the well-posedness of model (1) by showing two important properties, namely existence of solutions, and invariance of the number of vehicles along system trajectories. Then, we characterize the equilibria, we determine the minimum number of vehicles required for stabilizability (i.e., to ensure existence of equilibria), we show that without rebalancing vehicles the system is unstable (i.e., at some stations the number of waiting customers will grow without bounds), and, finally, we show how to optimally route the rebalancing vehicles across the transportation network (i.e., we discuss the optimal selection of each  $\alpha_{ij}$ ,  $i, j \in \mathcal{N}$ ).

#### A. Well-posedness of Model

The fluid model (1) is nonlinear, time-delayed, and the right-hand side is discontinuous. Due to the discontinuity, we need to analyze the model within the framework of Filippov solutions (see, e.g., [7]). The following proposition verifies the fluid model is well-posed.

*Proposition 3.1 (Well-posedness of fluid model):* For the fluid model (1), the following hold:

- (i) For every initial condition, there exist continuous functions  $c_i(t) : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$  and  $v_i(t) : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ ,  $i \in \mathcal{N}$ , satisfying the differential equations (1) in the Filippov sense.
- (ii) The total number of vehicles for  $t \geq 0$  is invariant and is equal to  $V = \sum_i v_i(0)$ .

*Proof:* To prove (i), it can be checked that all assumptions of Theorem II-1 in [10] for the existence of Filippov solutions to time-delayed differential equations with discontinuous right-hand side are satisfied, and the claim follows.

To prove (ii), let  $v_{ij}(t)$ , where  $t \geq 0$ , be the number of vehicles *in transit* from station  $i$  to station  $j$  (i.e., the vehicles for which the last station visited is  $i$  and the next station they will visit is  $j$ ). Clearly,  $v_{ii}(t) = 0$ . Now, the total number  $V(t)$  of vehicles in the system at time  $t$  is given by  $V(t) = \sum_{i=1}^n v_i(t) + \sum_{i,j} v_{ij}(t)$ . By definition,  $v_{ij}(t)$  can be expressed as

$$v_{ij}(t) = \int_{t-T_{ij}}^t p_{ij} \left( \lambda_i H(v_i(\tau)) - (\lambda_i - \mu_i) H(c_i(\tau)) H(v_i(\tau)) \right) + \alpha_{ij} H(v_i(\tau)) d\tau. \quad (2)$$

By applying the Leibniz integral rule, one can easily verify that  $\dot{V}(t) = \sum_{i=1}^n \dot{v}_i(t) + \sum_{i=1}^n \sum_{j=1}^n \dot{v}_{ij}(t) = 0$ , which proves the statement of the theorem. ■

#### B. Equilibria of Model

In order to determine the equilibria of our model we set  $\dot{c}_i = 0$  and  $\dot{v}_i = 0$  for all  $i \in \mathcal{N}$ . From the  $\dot{c}_i = 0$  equations we obtain

$$\lambda_i = \lambda_i H(v_i) - (\lambda_i - \mu_i) H(v_i) h(c_i). \quad (3)$$

Since  $\lambda_i < \mu_i$ , the above equations have a solution only if

$$c_i = 0 \quad \text{and} \quad v_i > 0 \quad \forall i \in \mathcal{N}.$$

Then, setting  $\dot{v}_i = 0$ , combined with (3), we obtain

$$0 = -\lambda_i + \sum_{j \neq i} \lambda_j p_{ji} - \gamma_i H(v_i) + \sum_{j \neq i} \alpha_{ji} H(v_j),$$

where we have used the fact that in a stationary equilibrium  $v_i(t)$  and  $c_i(t)$  are constants, independent of  $i$ .

For every equilibrium we require  $v_i > 0$ , and thus  $H(v_i) = 1$ . Therefore, if  $\alpha$  can be chosen such that

$$\sum_{j \neq i} \alpha_{ij} - \sum_{j \neq i} \alpha_{ji} = -\lambda_i + \sum_{j \neq i} \lambda_j p_{ji}, \quad (4)$$

for each  $i \in \mathcal{N}$ , then an equilibrium exists with  $c_i = 0$  and  $v_i > 0$  for each station  $i$ . Equation (4) shows that without

rebalancing vehicles (i.e., when each  $\alpha_{ij}$  is identically equal to zero), the system, in general, does not have equilibria. Furthermore, it can be shown that in absence of equilibria the system is, in general, unstable (the proof of this statement is rather straightforward—it can be obtained with a two station example—and is omitted in the interest of brevity). Hence, in general, *rebalancing vehicles are necessary to ensure equilibria and stability*. The next lemma shows that, fortunately, there always exists a *feasible rebalancing assignment*.

**Lemma 3.2 (Existence of feasible rebalancing assignment):** There always exists an assignment of  $\alpha$  such that equation (4) is satisfied.

*Proof:* Consider the assignment in which  $\alpha_{1k} := \lambda_k - \sum_{j \neq k} \lambda_j p_{jk}$ , for  $k \in \{2, \dots, n\}$ , and all other entries in  $\alpha$  are zero. For each  $i \in \{2, \dots, n\}$ , we can substitute the assignment into the LHS of (4), and verify that the constraint is satisfied. Thus, we just need to verify the assignment satisfies the constraint in (4) obtained when we set  $i = 1$ . Setting  $i = 1$  in (4), and bringing all terms to the LHS, the constraint becomes  $\sum_{k \neq 1} \alpha_{1k} + (\lambda_1 - \sum_{j \neq 1} \lambda_j p_{j1}) = 0$ . Substituting the assignment into the above equation, the LHS becomes  $\sum_k (\lambda_k - \sum_j \lambda_j p_{jk}) = \sum_k \lambda_k - \sum_k \sum_j \lambda_j p_{jk} = \sum_k \lambda_k - \sum_j \lambda_j \sum_k p_{jk} = 0$ , since  $\sum_k p_{jk} = 1$ . Thus, the proposed assignment is feasible and the linear system of equations given by (4) is consistent with at least one solution. ■

Let  $\mathcal{A}$  be the set of assignments  $\alpha$  such that equation (4) is satisfied. The next theorem gives a necessary and sufficient condition for the existence of equilibria.

**Theorem 3.3 (Existence of equilibria):** Model (1) admits equilibria if and only if the total number of vehicles  $V$  is larger than  $\underline{V} := \min_{\alpha \in \mathcal{A}} \sum_{ij} T_{ij} (p_{ij} \lambda_i + \alpha_{ij})$ .

*Proof:* Let us first consider sufficiency, and thus assume  $V > \underline{V}$ . Consider the feasible assignment  $\bar{\alpha} := \operatorname{argmin}_{\alpha \in \mathcal{A}} \sum_{ij} T_{ij} (p_{ij} \lambda_i + \alpha_{ij})$  (which exists by Lemma 3.2). When  $c_i = 0$  and  $v_i > 0$  for all  $i \in \mathcal{N}$ , the equilibrium number of transit vehicles is given by (recall equation (2)):  $\sum_{ij} T_{ij} (p_{ij} \lambda_i + \bar{\alpha}_{ij})$ . Hence, in order to satisfy the condition  $v_i > 0$  for all  $i \in \mathcal{N}$ , one needs a number of vehicles larger than  $\underline{V} := \sum_{ij} T_{ij} (p_{ij} \lambda_i + \bar{\alpha}_{ij})$ , which is verified by assumption. Necessity can be proven by similar arguments. ■ Theorem 3.3 can be easily generalized as follows.

**Corollary 3.4 (Existence for general assignments):**

Consider an assignment  $\alpha \in \mathcal{A}$ . If  $V > \sum_{ij} T_{ij} (p_{ij} \lambda_i + \alpha_{ij})$ , then the set of admissible equilibria is  $c_i = 0, v_i > 0, \forall i \in \mathcal{N}$ , where  $\sum_i v_i = V - \sum_{ij} T_{ij} (p_{ij} \lambda_i + \alpha_{ij})$ . If  $V \leq \sum_{ij} T_{ij} (p_{ij} \lambda_i + \alpha_{ij})$ , then no equilibrium exists.

### C. Optimal Rebalancing

Our objective is to minimize the average number of empty (rebalancing) vehicles traveling in the network, while ensuring stability. The time-average number of rebalancing vehicles is simply given by  $\sum_{i,j} T_{ij} \alpha_{ij}$ . Note that in minimizing this quantity, we are also minimizing the lower bound on the necessary number of vehicles  $\underline{V}$  from Theorem 3.3. Combining this objective with the existence of equilibria (i.e., the constraints

in (4), we see that  $\alpha$  should be chosen as the solution to the following minimum cost flow problem:

$$\begin{aligned} & \text{minimize} && \sum_{i,j} T_{ij} \alpha_{ij} \\ & \text{subject to} && \sum_{j \neq i} (\alpha_{ij} - \alpha_{ji}) = -\lambda_i + \sum_{j \neq i} \lambda_j p_{ji} && \forall i \in \mathcal{N} \\ & && \alpha_{ij} \geq 0 && \forall i, j \in \mathcal{N}, \end{aligned}$$

where  $\mathcal{N} := \{1, \dots, n\}$ . From Lemma 3.2, this linear program is feasible, and thus an optimal solution  $\alpha^*$  exists. The rebalancing policy is then given by sending empty vehicles from station  $i$  to station  $j$  at a rate of  $\alpha_{ij}^*$ .

When the travel times  $T_{ij}$  satisfy the triangle inequality, the above optimization can be simplified as follows. Let  $S$  be the set of stations with a *surplus* of vehicles, and  $D$  be the set of stations with a *deficit* vehicles; more formally, let  $S := \{i \in \mathcal{N} \mid \lambda_i < \sum_j \lambda_j p_{ji}\}$ , and  $D := \mathcal{N} \setminus S$ . By using the triangle inequality, one can easily show that in an optimal re-balancing solution, empty vehicles are sent from station  $i$  to station  $j$  (i.e.,  $\alpha_{ij}^* > 0$ ) only if  $i \in S$  and  $j \in D$ . This allows us to rephrase the above optimization as follows:

$$\begin{aligned} & \text{minimize} && \sum_{i \in S, j \in D} T_{ij} \alpha_{ij} \\ & \text{subject to} && \sum_{j \in D} \alpha_{ij} = -\lambda_i + \sum_{j \neq i} \lambda_j p_{ji} && \forall i \in S \\ & && \sum_{i \in S} \alpha_{ij} = \lambda_j - \sum_{i \neq j} \lambda_i p_{ij} && \forall j \in D \\ & && \alpha_{ij} \geq 0 && \forall i \in S, j \in D, \end{aligned}$$

This optimization is more computationally efficient than the first as it has fewer variables and fewer constraints.

## IV. ROBUST REBALANCING

In this section we consider the following questions: assume that the system is in equilibrium, what happens if there is a “burst” of customers arriving at the stations? Or, what happens if there is a sudden change in the number of available vehicles (e.g., because of a disruption)? In other words, we investigate the (local) stability of the equilibria of our model.

### A. Stability of Equilibrium Sets

We consider the following notion of local stability. Let  $\alpha \in \mathcal{A}$  be a feasible assignment, let  $V_\alpha := \sum_{ij} T_{ij} (p_{ij} \lambda_i + \alpha_{ij})$ , and assume  $V > V_\alpha$  (this is a necessary condition to have equilibria, see Corollary 3.4). The set of equilibria  $\mathcal{E}_\alpha := \{(\mathbf{c}, \mathbf{v}) \in \mathbb{R}^{2n} \mid c_i = 0, v_i > 0, \text{ for all } i \in \mathcal{N}, \text{ and } \sum_i v_i = V - V_\alpha\}$  is locally asymptotically stable if for any equilibrium  $(\underline{\mathbf{c}}, \underline{\mathbf{v}}) \in \mathcal{E}_\alpha$  there exists a neighborhood  $\mathcal{B}_\alpha^\delta(\underline{\mathbf{c}}, \underline{\mathbf{v}}) := \{(\mathbf{c}, \mathbf{v}) \in \mathbb{R}^{2n} \mid c_i \geq 0, v_i \geq 0 \text{ for all } i \in \mathcal{N}, \|\mathbf{c} - \underline{\mathbf{c}}, \mathbf{v} - \underline{\mathbf{v}}\| < \delta, \text{ and } \sum v_i = V - V_\alpha\}$  such that every evolution of model (1) starting at

$$\begin{aligned} c_i(\tau) &= 0 \text{ for } \tau \in [-\max_{i,j} T_{ij}, 0) \\ v_i(\tau) &= \underline{v}_i \text{ for } \tau \in [-\max_{i,j} T_{ij}, 0) \\ (\mathbf{c}(0), \mathbf{v}(0)) &\in \mathcal{B}_\alpha^\delta(\underline{\mathbf{c}}, \underline{\mathbf{v}}) \end{aligned} \tag{5}$$

has a limit which belongs to the equilibrium set. In other words,  $(\lim_{t \rightarrow +\infty} \mathbf{c}(t), \lim_{t \rightarrow +\infty} \mathbf{v}(t)) \in \mathcal{E}_\alpha$ .

The next theorem characterizes stability.

**Theorem 4.1 (Stability of equilibria):** Let  $\alpha \in \mathcal{A}$  be a feasible assignment, and assume  $V > V_\alpha$ ; then, the set of equilibria  $\mathcal{E}_\alpha$  is locally asymptotically stable.

*Proof:* Consider an equilibrium  $(\underline{\mathbf{c}}, \underline{\mathbf{v}}) \in \mathcal{E}_\alpha$  (note that  $\underline{\mathbf{c}} = \mathbf{0}$ ). We now prove that every evolution of model (1) starting at

$$\begin{aligned} c_i(\tau) &= 0 \text{ for } \tau \in [-\max_{i,j} T_{ij}, 0) \\ v_i(\tau) &= \underline{v}_i \text{ for } \tau \in [-\max_{i,j} T_{ij}, 0) \\ (\mathbf{c}(0), \mathbf{v}(0)) &\text{ such that } 0 \leq c_i(0) < v_i(0) \\ &\text{for all } i \in \mathcal{N}, \text{ and } \sum v_i(0) = V - V_\alpha, \end{aligned} \quad (6)$$

has a limit which belongs to the equilibrium set. The claim of the theorem will then be an easy consequence of this statement.

We start by observing the following fact. Assume that  $v_i(\tau) > 0$  for all  $\tau \in [-\max_{i,j} T_{ij}, t]$ , then at time  $t$  the differential equations read  $\dot{c}_i(t) = (\lambda_i - \mu_i)H(c_i(t))$ , for all  $i \in \mathcal{N}$ ; recalling that  $-\lambda_i + \sum_{j \neq i} \lambda_j p_{ji} - \gamma_i + \sum_{j \neq i} \alpha_{ji} = 0$ ,

$$\begin{aligned} \dot{v}_i(t) &= -\lambda_i + (\lambda_i - \mu_i)H(c_i) \\ &+ \sum_{j \neq i} p_{ji} \left( \lambda_j - (\lambda_j - \mu_j)H(c_j^i) \right) - \gamma_i + \sum_{j \neq i} \alpha_{ji} \\ &= (\lambda_i - \mu_i)H(c_i) - \sum_{j \neq i} p_{ji} (\lambda_j - \mu_j)H(c_j^i) \\ &\geq (\lambda_i - \mu_i)H(c_i), \quad \text{for all } i \in \mathcal{N}. \end{aligned}$$

Since  $v_i(\tau) > 0$  for all  $\tau \in [-\max_{i,j} T_{ij}, 0]$ , and since  $v_i(0) > c_i(0)$  for all  $i \in \mathcal{N}$ , we conclude that no  $v_i(t)$  can reach the value 0 *before* the corresponding number of customers  $c_i(t)$  has reached the value 0. However, once  $c_i(t)$  reaches the value 0 (after a time interval  $c_i(0)/(\mu_i - \lambda_i)$ ), the time derivative  $\dot{v}_i(t)$  is larger than or equal to zero. This implies that when the initial conditions satisfy (6), then  $v_i(t) > 0$  for all  $t \geq 0$ .

Since  $v_i(t) > 0$  for all  $t \geq 0$ , and since this implies that  $\dot{c}_i(t) = (\lambda_i - \mu_i)H(c_i(t))$  for all  $i \in \mathcal{N}$  and  $t \geq 0$ , we conclude that all  $c_i(t)$  will be equal to zero for all  $t \geq T' := \max_i c_i(0)/(\mu_i - \lambda_i)$ . Then, for  $t \geq T' + \max_{i,j} T_{ij} =: T''$  the differential equations become:  $\dot{c}_i(t) = 0$ ,  $\dot{v}_i(t) = 0$ .

Collecting the results obtained so far, we have that  $\lim_{t \rightarrow +\infty} c_i(t) = 0$  for all  $i \in \mathcal{N}$ . Moreover, since  $\dot{v}_i(t) = 0$  for all  $t \geq T''$ , the limit  $\lim_{t \rightarrow +\infty} v_i(t)$  exists. Finally, one has  $v_i(t) = v_i(0) + \int_0^t \dot{v}_i(\tau) d\tau \geq v_i(0) + \int_0^t \dot{c}_i(\tau) d\tau = v_i(0) + c_i(t) - c_i(0)$ . Since  $v_i(0) > c_i(0)$ , we conclude that  $\lim_{t \rightarrow +\infty} v_i(t) > 0$ . Thus any solution with initial conditions (6) has a limit which belongs to  $\mathcal{E}_\alpha$  (the property  $\lim_{t \rightarrow +\infty} \sum v_i(t) = V - V_\alpha$  is guaranteed by the invariance property in Proposition 3.1 and the assumption  $\sum v_i(0) = V - V_\alpha$ ).

Let  $\psi_i := v_i \sin \frac{\pi}{4}$  (see Figure 2), and let  $\psi_{\min} := \min_i \psi_i$ . Then, from the definitions of  $\psi_i$  and  $\psi_{\min}$ , it follows that if one chooses  $\delta = \psi_{\min}$ , then any solution of model (1) with

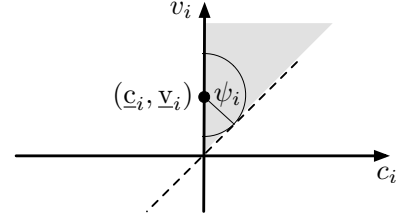


Fig. 2. The relation  $0 \leq c_i < v_i$ , and the definition of the radius  $\psi_i$ .

initial conditions satisfying (5) has a limit which belongs to the equilibrium set. This concludes the proof.  $\blacksquare$

## B. Stability of Single Equilibrium Points

So far, we have discussed the stability of the set of equilibria  $\mathcal{E}_\alpha$ , in the sense of equation (5). Here, we discuss stability of single equilibrium points. For simplicity, in this section we restrict perturbations to the number of customers only (the extension to the general case is straightforward).

Specifically, consider an assignment  $\alpha \in \mathcal{A}$ , let  $V_\alpha := \sum_{i,j} T_{ij} (p_{ij} \lambda_i + \alpha_{ij})$ , and assume, as usual,  $V > V_\alpha$ . We say that an equilibrium point  $(\underline{\mathbf{c}}, \underline{\mathbf{v}}) \in \mathcal{E}_\alpha$  is locally asymptotically stable if there exists a neighborhood

$$\mathcal{B}_\alpha^\delta(\underline{\mathbf{c}}, \underline{\mathbf{v}}) := \{(\mathbf{c}, \mathbf{v}) \in \mathbb{R}^{2n} \mid c_i \geq 0, v_i = \underline{v}_i \text{ for all } i \in \mathcal{N}, \text{ and } \|(\mathbf{c} - \underline{\mathbf{c}}, \mathbf{0})\| < \delta\}$$

such that every evolution of model (1) starting at the initial conditions in (5) satisfies  $\lim_{t \rightarrow +\infty} \mathbf{c}(t) = \mathbf{0}$  and  $\lim_{t \rightarrow +\infty} \mathbf{v}(t) = \underline{\mathbf{v}}$ . Note that we have redefined  $\mathcal{B}_\alpha^\delta(\underline{\mathbf{c}}, \underline{\mathbf{v}})$ .

One can show that, in general, single equilibrium points in  $\mathcal{E}_\alpha$  are *not* locally asymptotically stable; this implies that, to make individual equilibrium points locally asymptotically stable, an additional feedback term is needed. Specifically, assume that  $(\underline{\mathbf{0}}, \underline{\mathbf{v}}^d) \in \mathcal{E}_\alpha$  is a desired equilibrium point (e.g.,  $v_i^d = (V - V_\alpha)/n$ ). Then, one can make  $(\underline{\mathbf{0}}, \underline{\mathbf{v}}^d)$  locally asymptotically stable by adding to the vehicle's dynamics the feedback term

$$-H(v_i(t) - v_i^d) + \sum_{j \neq i} \frac{1}{n-1} H(v_j(t - T_{ji}) - v_j^d). \quad (7)$$

The normalizing constant  $1/(n-1)$  has the interpretation that the extra rebalancing vehicles are equally likely sent to the other stations. The next theorem shows that under model (1) *with* the additional feedback term (7) on vehicle's dynamics the equilibrium point  $(\underline{\mathbf{0}}, \underline{\mathbf{v}}^d)$  is locally asymptotically stable.

**Theorem 4.2 (Stability of single equilibrium points):** Let  $\alpha \in \mathcal{A}$  be a feasible assignment, assume  $V > V_\alpha$ , and let  $(\underline{\mathbf{0}}, \underline{\mathbf{v}}^d) \in \mathcal{E}_\alpha$  be a desired equilibrium point. Under model (1), *with* the additional feedback term (7) on the vehicle's dynamics, the equilibrium point  $(\underline{\mathbf{0}}, \underline{\mathbf{v}}^d)$  is locally asymptotically stable in the sense of equation (5).

*Proof:* The proof is similar to the proof of Theorem 4.1. Assume perturbations in the number of customers  $\underline{\mathbf{c}}$  such that  $0 \leq c_i(0) < v_i^d = v_i(0)$  for all  $i \in \mathcal{N}$ .

We start by observing the following fact. Assume that  $v_i(\tau) > 0$  for all  $\tau \in [-\max_{i,j} T_{ij}, t]$ , then at time  $t$  the

differential equations read  $\dot{c}_i(t) = (\lambda_i - \mu_i)H(c_i(t))$ , for all  $i \in \mathcal{N}$ ; recalling that  $-\lambda_i + \sum_{j \neq i} \lambda_j p_{ji} - \gamma_i + \sum_{j \neq i} \alpha_{ji} = 0$ , one obtains

$$\begin{aligned} \dot{v}_i &= -\lambda_i + (\lambda_i - \mu_i)H(c_i) \\ &+ \sum_{j \neq i} p_{ji} \left( \lambda_j - (\lambda_j - \mu_j)H(c_j^i) \right) - \gamma_i + \sum_{j \neq i} \alpha_{ji} \\ &- H(v_i - v_i^d) + \sum_{j \neq i} \frac{1}{n-1} H(v_j^i - v_i^d) \\ &= (\lambda_i - \mu_i)H(c_i) - \sum_{j \neq i} p_{ji} (\lambda_j - \mu_j)H(c_j^i) \\ &- H(v_i - v_i^d) + \sum_{j \neq i} \frac{1}{n-1} H(v_j^i - v_i^d) \\ &\geq (\lambda_i - \mu_i)H(c_i) + \begin{cases} -1 & \text{if } v_i > v_i^d, \\ 0 & \text{if } v_i \leq v_i^d, \end{cases} \end{aligned}$$

for all  $i \in \mathcal{N}$ . Since  $v_i(\tau) > 0$  for all  $\tau \in [-\max_{i,j} T_{ij}, 0]$ , and since  $v_i(0) = v_i^d > c_i(0)$  for all  $i \in \mathcal{N}$ , we conclude that no  $v_i(t)$  can reach the value 0 *before* the corresponding number of customers  $c_i(t)$  has reached the value 0. However, once  $c_i(t)$  reaches the value 0 (after a time interval  $c_i(0)/(\mu_i - \lambda_i)$ ), the time derivative  $\dot{v}_i(t)$  is larger than or equal to zero whenever  $v_i(t) < v_i^d$ . This implies that under the assumptions on the initial conditions, then  $v_i(t) > 0$  for all  $t \geq 0$ .

Since  $v_i(t) > 0$  for all  $t \geq 0$ , and since this implies that  $\dot{c}_i(t) = (\lambda_i - \mu_i)H(c_i(t))$  for all  $i \in \mathcal{N}$  and  $t \geq 0$ , we conclude that all  $c_i(t)$  will be equal to zero for all  $t \geq T' := \max_i c_i(0)/(\mu_i - \lambda_i)$ . Then, for  $t \geq T' + \max_{i,j} T_{ij} =: T''$  the differential equations become:

$$\begin{aligned} \dot{c}_i(t) &= 0, \\ \dot{v}_i(t) &= -H(v_i(t) - v_i^d) + \sum_{j \neq i} \frac{1}{n-1} H(v_j(t - T_{ji}) - v_i^d). \end{aligned}$$

Note that  $\dot{v}_i(t) \leq 0$  if  $v_i(t) > v_i^d$  and  $\dot{v}_i(t) \geq 0$  if  $v_i(t) < v_i^d$ . Hence, if at some instant  $t \geq T''$  the number of vehicles satisfies  $v_i(t) = v_i^d$ , then  $v_i(\tau) = v_i^d$  for all  $\tau \geq t$ , in other words  $v_i$  will “slide” along the mode  $v_i^d$  for ever. Let  $\mathcal{U}(t)$  be the set of indexes of stations such that  $v_i(t) > v_i^d$ , and  $\mathcal{D}(t)$  be the set of indexes of stations such that  $v_i(t) < v_i^d$ .

One can show that  $\mathcal{U}(t) \neq \emptyset$  if and only if  $\mathcal{D}(t) \neq \emptyset$ , for all  $t \geq T''$ . Indeed, assume by the sake of contradiction, that  $\mathcal{U}(t) \neq \emptyset \wedge \mathcal{D} = \emptyset$ . Then, since  $\mathcal{D}(t) = \emptyset$ , and since  $c_i = 0$  and  $v_i > 0$  for all  $i \in \mathcal{N}$  and for all  $t \geq T'$ , one can write for all  $t \geq T''$ :  $V(t) = \sum_i^n v_i(t) + \sum_{i,j} v_{ij}(t) > \sum_i v_i^d + \sum_{i,j} T_{ij}(p_{ij}\lambda_i + \bar{\alpha}_{ij}) = V(0)$ , which is in contradiction with the fact the the total number of vehicles is invariant. The proof for the converse implication is identical.

Now, if  $\mathcal{U}(t) = \emptyset$ , then one immediately obtains the claim of the theorem. Assume, instead, that  $\mathcal{U}(t) \neq \emptyset$ ; from the previous discussion, one concludes that for each  $i \in \mathcal{U}(t)$ ,  $\dot{v}_i(t) \leq -1/(n-1)$ , since  $\mathcal{D}(t) \neq \emptyset$ . This implies that for  $t \geq T'' + (n-1) \max_{i \in \mathcal{U}(T'')} v_i(T'')$  the set  $\mathcal{U}(t)$  will be empty, and thus also set  $\mathcal{D}(t)$  will be empty, implying that  $v_i = v_i^d$  from then on. This concludes the proof. ■

## V. ADAPTIVE REAL-TIME REBALANCING POLICY

Until now, the policies have required knowledge of the arrival rates  $\lambda_{ij}$ , and the destination distribution  $p_{ij}$ . In this section we introduce a policy that does not require any *a priori* information. The idea is to repeatedly solve the optimization introduced in Section III-C, but using the current distribution of customers and vehicles. Let us define the number of vehicles *owned* by a station  $i$  to be the number of vehicles at that station, given by  $v_i(t)$ , plus the number of vehicles in transit to the station, given by  $\sum_j v_{ji}(t)$ :  $v_i^{\text{own}}(t) := v_i(t) + \sum_j v_{ji}(t)$ . Note that by definition,  $\sum_i v_i^{\text{own}}(t) = V$  at all times  $t \geq 0$ .

Now, suppose that our objective is to have an equal number of vehicles at each station. If station  $i$  has  $c_i(t)$  customers and  $v_i(t)$  vehicles, then  $\min\{c_i(t), v_i(t)\}$  vehicles will leave the station to serve the waiting customers. The excess vehicles at station  $i$  is then  $v_i^{\text{excess}}(t) := v_i^{\text{own}}(t) - c_i(t)$ . These are the vehicles that station  $i$  currently has available to send to other stations in need. Thus, the total number of excess vehicles in the system is  $\sum_i v_i^{\text{excess}}(t) = V - \sum_i c_i(t)$ . Note that this number may be negative. We would like to split these excess vehicles among the  $n$  stations according to some desired distribution. That is, for each station  $i$  we have a desired number of vehicles  $v_i^d(t)$ , such that  $\sum_i v_i^d(t) \leq V - \sum_i c_i(t)$  for all  $t$ . As an example, we may choose that the excess vehicles are evenly split among the stations so that

$$v_i^d(t) = \left\lfloor \frac{V - \sum_j c_j(t)}{n} \right\rfloor \quad \text{for each station } i, \quad (8)$$

where we take the floor to obtain an integer. Our goal is to have  $v_i^{\text{excess}}(t) \geq v_i^d(t)$  for all  $t$ .

Let us define an optimization horizon  $t_{\text{hor}} > 0$ . At time instants  $kt_{\text{hor}}$ , where  $k$  is a non-negative integer, we rebalance the excess vehicles by solving the following optimization:

$$\begin{aligned} &\text{minimize} \quad \sum_{i,j} T_{ij} \text{num}_{ij} \\ &\text{subject to} \quad v_i^{\text{excess}}(t) + \sum_{j \neq i} (\text{num}_{ji} - \text{num}_{ij}) \geq v_i^d(t) \quad \forall i \in \mathcal{N}, \\ &\quad \text{num}_{ij} \in \mathbb{N} \quad \quad \quad \forall i, j \in \mathcal{N}, i \neq j, \end{aligned}$$

Note that this is an integer linear program, where  $\text{num}_{ij}$  is the number of vehicles that station  $i$  will send to station  $j$ . It can be written in the form  $\{\min cx \mid Ax \geq b, x \in \mathbb{N}\}$ . However, it can easily be verified that the constraint matrix  $A$  is *totally unimodular* [11]. In addition, the vector  $b$  contains integer entries  $v_i^{\text{excess}}(t) - v_i^d(t)$ . Therefore, we can relax the integer constraints to  $\text{num}_{ij} \geq 0$ , and solve the problem as a linear program  $\{\min cx \mid Ax \geq b, x \geq 0\}$ . The resulting solution will necessarily have integer values, and thus will also be the optimal solution to the integer linear program.

Therefore, we can efficiently rebalance the system every  $t_{\text{hor}}$  time units without knowledge of  $\lambda_i$  or  $p_{ij}$ . Each time the optimization is solved, we simply send  $\text{num}_{ij}$  rebalancing vehicles from station  $i$  to station  $j$ . In the next section we will characterize the performance of this policy in simulation. For future work, we are looking at modelling customer arrivals

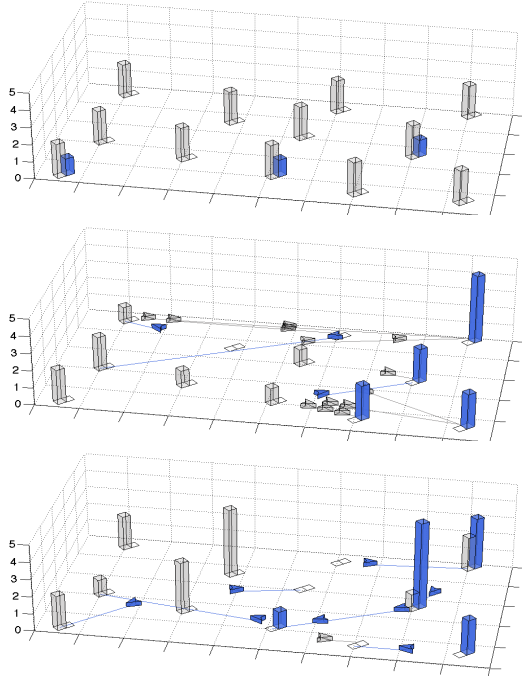


Fig. 3. The simulation environment for load balancing in mobility-on-demand systems. At each station, the grey bar shows the number of vehicles, and the blue bar shows the number of waiting customers. Blue vehicles are carrying customers, while grey vehicles are performing rebalancing trips. The first frame shows the initial condition, with 24 vehicles and 3 customers. The second and third frames show the system as it evolves.

as a stochastic process and then analyzing the theoretical performance of this policy.

## VI. SIMULATIONS

We have developed a simulation environment in MATLAB<sup>®</sup> for testing rebalancing policies. An example of the environment is shown in Figure 3. In this environment customer arrive *stochastically* at each station  $i$  according to a Poisson process with parameter  $\lambda_i$ . Each customer's destination is sampled from the distribution  $p_{ij}$ . Since the evolution of the system is stochastic, we perform several trials, and then compute statistics in order to characterize a policy's performance. For each policy, we solve the necessary linear programs using the freely available SeDuMi (Self-Dual-Minimization) toolbox. Simulations were run on a laptop computer with a 2.66 GHz dual core processor and 4 GB of RAM.

This section presents simulation results for the 12 station environment in Figure 3. We have also performed simulations for larger scale environments containing  $> 50$  stations. The travel times  $T_{ij}$  between stations is given by the Euclidean distance. The dimensions of the environment are 10 by 20 (dimensionless) units, and each vehicle moves 0.2 units per time step. The arrival rates at each station were randomly selected, as was the destination density  $p_{ij}$ . Using the necessary condition on the number of vehicles in (see Theorem 3.3), we obtain that  $V \geq 13.4$  for any stabilizing policy.

### A. Real-time Rebalancing Policy

Figure 4 summarizes the performance of the real-time rebalancing policy of Section V with  $v_i^d(t)$  as shown in (8). The left figure shows the number of waiting customers, and the number of in-transit vehicles as a function of the optimization horizon  $t_{\text{hor}}$ . The total number of vehicles is  $V = 36$ . Each data point is the mean of 20 independent trials, where each trial starts from an initial condition of equally distributed vehicles, and no customers, and runs for 5000 time steps. Error bars show the standard deviation over the 20 trials. We see that as the optimization horizon increases, the number of in-transit vehicles decreases, but the number of waiting customers increases. Thus, the optimal choice of  $t_{\text{hor}}$  is a trade-off between the cost of performing rebalancing trips, and the wait-time of customers.

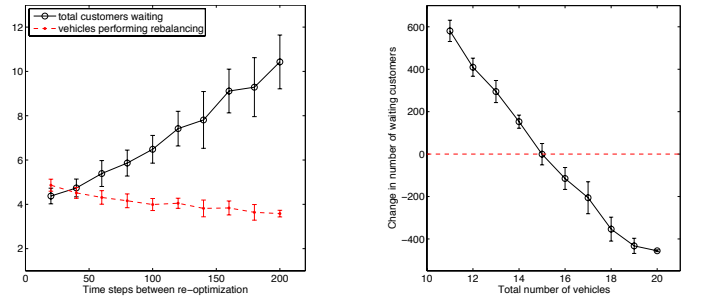


Fig. 4. Analysis of the real-time rebalancing policy. Left figure: The number of customers or vehicles as a function of the optimization horizon  $t_{\text{hor}}$ . Right figure: Stability as a function of the number of vehicles  $V$ . Positive numbers indicate instability, while negative numbers indicate stability.

The right figure shows the stability of the real-time rebalancing policy. Each data point shows the mean of 20 trials. In each trials, we start the system with 480 customers, and run the system for 15,000 time steps. We then look at the difference between the initial number of customers, and the time-average present over the last 1000 time steps of the trial. From the simulation, we can see that 15 vehicles is the threshold for stability in this example. With  $V \geq 15$  we have stability. This compares quite well with the necessary condition of  $V > 13.4$ .

### B. Policy Comparison

Finally, we compare the performance of three policies: 1) the basic policy from the fluid model, where we send a constant rate of  $\alpha_{ij}$  vehicles between stations, 2) the fluid model policy with feedback, where the desired number of vehicles is set to  $v^d = \lceil V/n \rceil$  and 3) the real-time rebalancing policy with  $t_{\text{hor}} = 100$  and  $v_i^d(t)$  as shown in (8). For each policy, Figure 5 (left) shows the number of waiting customers as a function of the total number of vehicles  $V$ , on a log scale. Figure 5 (right) shows the number of in-transit vehicles as a function of  $V$ . Each data point is the mean of 20 independent trials, and each trial consists of 5000 time steps. In each trial we compute mean number of waiting customers, and the mean number of in-transit vehicles over the last 2000 time steps of the trial. Thus, a finite number of waiting vehicles does not necessarily indicate stability. We can see from Figure 4 that at least 15 vehicles are needed for stability.

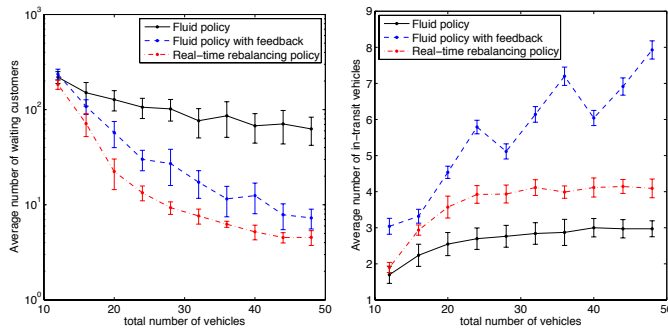


Fig. 5. The performance of the three different policies. Left figure: The time-average number of customers waiting in the system. Right figure: The time-average number of vehicles performing rebalancing trips per unit of time.

From the left figure we see that the real-time rebalancing policy has the best performance in terms of number of waiting customers. The basic fluid policy performs quite poorly. This is due to the stochastic fluctuations in the customer arrival rates, and their chosen destinations. The fluid policy with feedback performs adequately, but consistently has over 50% more waiting customers than the real-time rebalancing policy.

From the right figure we see that the fluid model policy sends the fewest rebalancing vehicles. Thus, to minimize rebalance cost without regard to customer satisfaction (i.e., wait times), the fluid model policy performs best. The real-time rebalancing policy sends fewer vehicles than the fluid policy with feedback. Thus, the real-time policy outperforms the feedback policy in terms of waiting customers, and in-transit vehicles. Finally, the number of in-transit vehicles for the fluid policy with feedback increases in a sawtooth pattern. The peaks occur at 24, 36, and 48 vehicles. This occurs because the desired number of vehicles at a station was chosen as  $v^d = \lceil V/n \rceil$ . Since there are  $n = 12$  stations, when the number of vehicles is increased from 24 to 25, the value  $v^d$  changes from 2 to 3. This shift dramatically affects the number of in-transit vehicles. However, we can see that it does not have as large of an effect on the number of waiting customers. Thus, one may be able to improve the performance of the fluid policy with feedback by setting  $v^d$  to a value larger than  $\lceil V/n \rceil$ .

## VII. CONCLUSION

In this paper we studied the problem of rebalancing a mobility-on-demand system to efficiently transport customers in an urban environment. We proposed a robotic solution to rebalancing that involves empty robotic vehicles autonomously driving between stations. For a fluid model of the system, we showed that the optimal rebalancing policy can be found as the solution to a linear program. Under this policy, every station reaches an equilibrium in which there are excess vehicles and no waiting customers. We used this solution to develop a real-time rebalancing policy that can operate under stochastic customer demand. For future work we plan to analyze the stochastic queueing model and characterize the performance of the real-time rebalancing policy. We also plan to enrich our model by including uncertainty in the travel times and

time windows for the customers. Finally, we are interested in using dynamic pricing to provide incentives for customers to perform rebalancing trips themselves.

## ACKNOWLEDGEMENTS

This research was initiated while M. Pavone and S. L. Smith were at the Massachusetts Institute of Technology. This research was supported in part by the Future Urban Mobility project of the Singapore-MIT Alliance for Research and Technology (SMART) Center, with funding from Singapore's National Research Foundation. It was also supported by ONR grant N000140911051 and NSF EFRI0735953.

## REFERENCES

- [1] R. Balakrishna, M. Ben-Akiva, and H. N. Koutsopoulos. Offline Calibration of Dynamic Traffic Assignment: Simultaneous Demand-and-Supply Estimation. *Transportation Research Record: Journal of the Transportation Research Board*, 2003(1):pp. 50–58, 2007.
- [2] G. Berbeglia, J. F. Cordeau, and G. Laporte. Dynamic pickup and delivery problems. *European Journal of Operational Research*, 202(1): 8 – 15, 2010.
- [3] M. Buehler, K. Iagnemma, and S. Singh, editors. *The 2005 DARPA Grand Challenge: the Great Robot Race*. Number 36 in Springer Tracts in Advanced Robotics. Springer, 2007.
- [4] M. Buehler, K. Iagnemma, and S. Singh, editors. *The DARPA Urban Challenge: Autonomous Vehicles in City Traffic*. Number 56 in Springer Tracts in Advanced Robotics. Springer, 2009.
- [5] V. Cardellini, M. Colajanni, and P. S. Yu. Dynamic load balancing on Web-server systems. *IEEE Internet Computing*, 3(3):28 – 39, 1999.
- [6] George Cybenko. Dynamic load balancing for distributed memory multiprocessors. *Journal of Parallel and Distributed Computing*, 7(2): 279 – 301, 1989.
- [7] A. F. Filippov. *Differential Equations with Discontinuous Righthand Sides*, volume 18 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1988.
- [8] T. L. Friesz, J. Luque, R. L. Tobin, and B. W. Wie. Dynamic Network Traffic Assignment Considered as a Continuous Time Optimal Control Problem. *Operations Research*, 37(6):pp. 893–901, 1989.
- [9] T. L. Friesz, D. Bernstein, T. E. Smith, R. L. Tobin, and B. W. Wie. A Variational Inequality Formulation of the Dynamic Network User Equilibrium Problem. *Operations Research*, 41(1):pp. 179–191, 1993.
- [10] G. Haddad. Monotone trajectories of differential inclusions and functional differential inclusions with memory. *Israel journal of mathematics*, 39(1-2):83 – 100, 1981.
- [11] B. Korte and J. Vygen. *Combinatorial Optimization: Theory and Algorithms*, volume 21 of *Algorithmics and Combinatorics*. Springer, 4 edition, 2007.
- [12] D. K. Merchant and G. L. Nemhauser. Optimality Conditions for a Dynamic Traffic Assignment Model. *Transportation Science*, 12(3):200–207, 1978.
- [13] P. Midgley. The role of smart bike-sharing systems in urban mobility. Technical report, 2009.
- [14] W. J. Mitchell, C. E. Borroni-Bird, and L. D. Burns. *Reinventing the Automobile: Personal Urban Mobility for the 21st Century*. The MIT Press, Cambridge, MA, 2010.
- [15] United Nations. World Urbanization Prospects: The 2007 Revision Population Database. Technical report, 2007.
- [16] Bureau of Transportation Statistics. Highlights of the 2001 National Household Travel Survey. Technical report, 2001.
- [17] S. N. Parragh, K. F. Doerner, and R. F. Hartl. A survey on pickup and delivery problems. *Journal fur Betriebswirtschaft*, 58(2):81 – 117, 2008.
- [18] M. Pavone, K. Treleaven, and E. Frazzoli. Fundamental Performance Limits and Efficient Polices for Transportation-On-Demand Systems. In *IEEE Conf. Decision and Control*, pages 5622–5629, Atlanta, GA, 2010.
- [19] S. Peeta and A. Ziliaskopoulos. Foundations of Dynamic Traffic Assignment: The Past, the Present and the Future. *Networks and Spatial Economics*, 1:233–265, 2001.
- [20] A. K. Ziliaskopoulos. A Linear Programming Model for the Single Destination System Optimum Dynamic Traffic Assignment Problem. *Transportation Science*, 34(1):37–49, 2000.