Intermodal Autonomous Mobility-on-Demand

Mauro Salazar[1,2], Nicolas Lanzetti[1,2], Federico Rossi[2], Maximilian Schiffer[2,3], and Marco Pavone[2]

***Abstract*—In this paper we study models and coordination policies for intermodal Autonomous Mobility-on-Demand (AMoD), wherein a fleet of self-driving vehicles provides on-demand mobility jointly with public transit. Specifically, we first present a network flow model for intermodal AMoD, where we capture the coupling between AMoD and public transit and the goal is to maximize social welfare. Second, leveraging such a model, we design a pricing and tolling scheme that allows the system to recover a social optimum under the assumption of a perfect market with selfish agents. Third, we present real-world case studies for the transportation networks of New York City and Berlin, which allow us to quantify the general benefits of intermodal AMoD, as well as the societal impact of different vehicles. In particular, we show that vehicle size and powertrain type heavily affect intermodal routing decisions and, thus, system efficiency. Our studies reveal that the cooperation between AMoD fleets and public transit can yield significant benefits compared to an AMoD system operating in isolation, whilst our proposed tolling policies appear to be in line with recent discussions for the case of New York City.**

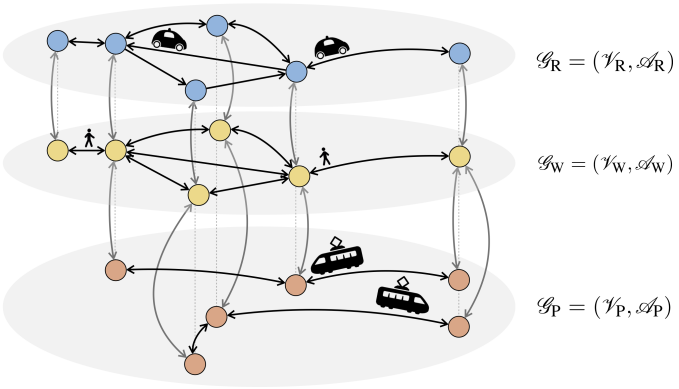

Fig. 1. The I-AMoD network consists of a road digraph, a walking digraph and a public transportation digraph. The colored dots denote intersections or stops and the black arrows represent road links, pedestrian pathways or public transit arcs. The grey dotted lines denote geographically close nodes while the grey arrows are the mode-switching arcs connecting them.

I. Introduction

Traffic congestion is soaring all around the world. Besides mere discomfort for passengers, congestion causes severe economic and environmental harm, e.g., due to the loss of working hours and pollutant emissions such as CO_2, particulate matter, and NO_x [1]. In 2013, traffic congestion cost U.S. citizens 124 Billion USD [2]. Notably, transportation remains one of a few sectors in which emissions are still increasing [3]. Governments and municipalities are struggling to find sustainable ways of transportation that can match mobility needs and reduce environmental harm as well as congestion.

To achieve sustainable modes of transportation, new mobility concepts and technology changes are necessary. However, the potential to realize such concepts in urban environments is limited, since upgrades to available infrastructures (e.g., roads and subway lines) and their capacity are often extremely costly and require decades-long planning timelines. Thus, mobility concepts that use existing infrastructure in a more efficient way are especially attractive. In this course, mobility-on-demand services appear to be particularly promising. Herein, two main concepts exist. On the one hand, free floating car sharing systems strive to reduce the total number of private vehicles in city centers. However, these systems offer limited flexibility and are generally characterized by low adoption rates that result from low vehicle availabilities due to the difficulty of *rebalancing* empty vehicles to counter asymmetric customer demand [4], [5]. On the other hand, ride-hailing systems aim to enhance and extend the service of taxi fleets. However, current studies show that ride-hailing services can worsen traffic congestion significantly due to the induced demand and the vehicle-miles traveled by empty vehicles. Moreover, additional demand may also result from a shift in the modal split as ride-hailing operators offer a low cost point-to-point connection. Indeed, recent studies for the Manhattan area revealed the massive magnitude of such effects: from 2013 to 2018, the number of for-hire vehicles exploded from 47,000 to 103,000, 68,000 of which are employed for ride-hailing services. Due to this increase, the average traffic speed dropped to 4.7 mph, which equals a brisk walk [6].

Autonomous Mobility-on-Demand (AMoD) systems hold promise as a future mobility concept in urban environments. They comprise a fleet of robotic, self-driving vehicles that transport passengers between their origins and destinations. A central operator runs such systems by assigning passenger requests to vehicles and computing rebalancing routes for the unassigned empty vehicles, in order to re-align their geographical distribution with demand for transportation. Thus, an AMoD system can replace a conventional taxi, car sharing, or ride-hailing fleet, while offering several advantages compared to the previously discussed concepts: First, no relocation costs for drivers arise; second, due to continuous rebalancing, much higher vehicle utilization rates can be achieved; third, the centralized control of the complete fleet allows for more operational flexibility and efficiency compared to ride-hailing or taxi fleets where a central operator can suggest, but not directly control, vehicle routes. This enables the operator to adopt global fleet-wide routing strategies to mitigate congestion.

[1]Institute for Dynamic Systems and Control ETH Zürich, Zurich (ZH), Switzerland {samauro,lnicolas}@ethz.ch

[2]Department of Aeronautics and Astronautics, Stanford University, Stanford (CA), United States {samauro,frossi2,pavone}@stanford.edu

[3]TUM School of Management, Technical University of Munich, 80333 Munich, Germany schiffer@tum.de

However, despite these benefits, AMoD systems operating in isolation might still worsen congestion due to shifts in the modal share. To secure sustainable and congestion-free urban mobility, an AMoD system should rather interact with and complement existing mass transit options. Against this backdrop, our study develops modeling and optimization tools to assess the benefits of an intermodal transportation system that combines public transit with AMoD (Fig. 1).

Related literature: Our work contributes to three different research streams, namely: AMoD systems, congestion pricing, and multimodal passenger transport, which we review in the following.

A number of approaches to characterize and control AMoD systems in isolation are available, ranging from queuing-theoretical models [7]–[9] to simulation-based models [10]–[12] and multi-commodity network flow models [13]–[15]. Queueing-theoretical models capture the stochasticity of the customer arrival process and are amenable to efficient control synthesis. However, their complex structure makes it difficult to capture the interaction with other modes of transportation. Simulation-based models capture transportation systems with very high fidelity, incorporating complex choice models and microscopic interactions, but are generally not amenable to efficient optimization. Network flow models are amenable to efficient optimization and allow for the inclusion of a variety of complex constraints. Accordingly, they have seen wide use in problems ranging from control of AMoD systems in congested road networks [13], [16], [17], to cooperative control of AMoD systems and the electric power network [18], and control of human-operated MoD systems [19].

Congestion pricing in general has been widely investigated, and a body of theoretical work [20]–[23] and experimental results [24]–[26] are available. However, only few approaches focus on pricing in the context of AMoD: Specifically, [27] focuses on congestion pricing for self-driving vehicles by incentivizing socially and environmentally aware travel modes, while [28] proposes pricing schemes to foster the use of AMoD systems. However, these studies comprise logit modeling approaches and rely on agent-based simulations that assess the performance of pre-determined intermodal AMoD (I-AMoD) routing policies. In contrast, our optimization-based approach identifies the *best* achievable performance of an I-AMoD system and enables the synthesis of policies that steer a system towards such an optimum.

Literature on intermodal passenger transportation including Mobility-on-Demand (MoD) and AMoD is still sparse. First, studies on the interplay between AMoD and public transportation exist, focusing either on fluidic [29] or simulation-based [11], [30], [31] models. However, these studies focus on the analysis of specific scenarios, as opposed to the *optimization* of joint control policies for AMoD systems and public transit. In general, to the best of the authors' knowledge, only descriptive analyses of intermodal passenger transport including MoD exist [32].

In summary, some optimization approaches and control policies for AMoD systems are available. These approaches, however, do not capture the interaction between AMoD and public transit. Focusing on pricing schemes, existing studies address individual externalities (e.g., congestion), but no study captures the interplay between multiple externalities arising from the synchronization of different modes of transportation. To date, there exist no optimization frameworks that capture optimal coordination policies for I-AMoD systems whilst assessing their achievable performance.

Statement of contributions: The goal of this paper is to introduce a mesoscopic optimization approach for I-AMoD systems. Specifically, the contribution of this paper is fourfold: First, we develop a multi-commodity network flow optimization model that captures the joint operations of AMoD systems and public transit. In our model, the objective is to maximize the social welfare, i.e., to minimize the customers' travel time together with the operational costs of different transportation modes. Herein, we also consider energy consumption, pollution, and congestion effects. Second, we propose a pricing and tolling scheme that helps to realize the social optimum in the presence of selfish customers and AMoD operators. Third, we present real-world case studies for New York City (NYC) and Berlin accounting for the impact of the urban transportation network and of the AMoD vehicles' characteristics on the achievable societal costs, including travel times and emissions. Fourth, we derive managerial insights on the benefits of I-AMoD systems: our results show that an I-AMoD system can significantly reduce travel times, pollutant emissions, total number of cars, and overall costs compared to an AMoD system operating in isolation. Interestingly, our pricing and tolling scheme is aligned with recently proposed congestion surcharges for ride-hailing vehicles [33].

A preliminary version of this paper was presented at the 2018 Intelligent Transportation Systems Conference [34]. In this revised and extended version, we broaden the discussion of the literature, detail models for travel time, road congestion, and energy consumption, present a rigorous optimality proof for the proposed pricing scheme, and discuss new numerical results for the central neighborhoods of NYC and Berlin.

Organization: The remainder of this paper is structured as follows: In Section II we present a flow optimization model for I-AMoD. Section III derives a pricing and tolling scheme to steer self-interested agents towards a social optimum. Section IV presents case studies for NYC and Berlin, which are complementary in terms of spatial displacement, road network structure, and public transit. Finally, Section V concludes the paper with a short summary and an outlook on future research.

II. Network Flow Model for I-AMoD

This section presents a network flow optimization approach for intermodal AMoD. In particular, we consider 1) the assignment of customer requests to transport flows, 2) different modes of transportation, 3) road capacity limits, 4) vehicle based energy consumption models for the AMoD fleet, and 5) rebalancing operations to reposition empty vehicles according to mobility demand. Assuming a centrally controlled system, we introduce a network flow model in Section II-A, and discuss the representation of travel time and road congestion in Section II-B. Section II-C presents our energy consumption model. Section II-D details the I-AMoD optimization problem

and its objective. Finally, Section II-E discusses our modeling assumptions. Readers not familiar with basic elements from graph theory are referred to Appendix A.

A. Multi-commodity Flow Model

We model the transportation system and its different transportation modes on a digraph $\mathscr{G} = (\mathscr{V}, \mathscr{A})$ as shown in Fig. 1. It consists of a set of nodes $\mathscr{V}$ and a set of arcs $\mathscr{A} \subseteq \mathscr{V} \times \mathscr{V}$, containing a road network layer $\mathscr{G}_\mathrm{R} = (\mathscr{V}_\mathrm{R}, \mathscr{A}_\mathrm{R})$, a public transportation layer $\mathscr{G}_\mathrm{P} = (\mathscr{V}_\mathrm{P}, \mathscr{A}_\mathrm{P})$, and a walking layer $\mathscr{G}_\mathrm{W} = (\mathscr{V}_\mathrm{W}, \mathscr{A}_\mathrm{W})$. The road network consists of intersections $i \in \mathscr{V}_\mathrm{R}$ and road links $(i,j) \in \mathscr{A}_\mathrm{R}$. We model public transportation, i.e., tram and subway lines as distinct trees, using a set of station nodes $i \in \mathscr{V}_\mathrm{P}$ and a set of line segments $(i,j) \in \mathscr{A}_\mathrm{P}$. The walking network comprises walkable streets $(i,j) \in \mathscr{A}_\mathrm{W}$ between intersections $i \in \mathscr{V}_\mathrm{W}$. Finally, we model the possibility of customers switching transportation modes (e.g., exiting the subway or hailing an AMoD ride) by connecting the pedestrian layer to the road and public transportation layers with a set of mode-switching arcs $\mathscr{A}_\mathrm{C} \subseteq \mathscr{V}_\mathrm{R} \times \mathscr{V}_\mathrm{W} \cup \mathscr{V}_\mathrm{P} \times \mathscr{V}_\mathrm{W}$, whereby $\mathscr{V}_\mathrm{R} \cap \mathscr{V}_\mathrm{P} = \emptyset$. Accordingly, $\mathscr{V} = \mathscr{V}_\mathrm{W} \cup \mathscr{V}_\mathrm{R} \cup \mathscr{V}_\mathrm{P}$ and $\mathscr{A} = \mathscr{A}_\mathrm{W} \cup \mathscr{A}_\mathrm{R} \cup \mathscr{A}_\mathrm{P} \cup \mathscr{A}_\mathrm{C}$ hold. Given the structural properties of road and walking networks in urban environments, we make the following assumption without loss of generality.

Assumption 1. *The graphs $\mathscr{G}$, $\mathscr{G}_\mathrm{R}$, and $\mathscr{G}_\mathrm{W}$ are strongly connected.*

Traversing an arc (i,j) of length s_{ij} takes on average t_{ij} time units. For mode-switching arcs, t_{ij} denotes the time needed to switch between two means of transportation. We represent travel requests as follows.

Definition II.1 (Requests)**.** *A request r is a triple $(o,d,\alpha) \in \mathscr{V} \times \mathscr{V} \times \mathbb{R}_+$, given by its origin node o, its destination node d, and its request rate $\alpha > 0$. We denote a set of M requests by $\mathscr{R} \coloneqq \{r_m\}_{m \in \mathscr{M}}$, where $\mathscr{M} \coloneqq \{1, \dots, M\}$.*

Assumption 2. *All requests appear on the walking digraph, i.e., $o_m, d_m \in \mathscr{V}_\mathrm{W},\ \forall m \in \mathscr{M}$.*

Considering the different transportation modes, $f_m(i,j)$ denotes the flow (i.e., the number of customers per unit time) traversing arc $(i,j) \in \mathscr{A}$ for a certain travel request m. To account for rebalancing flows of AMoD vehicles between a customer's drop-off and the next customer's pick-up, $f_0(i,j)$ denotes the flow of empty vehicles on road arcs $(i,j) \in \mathscr{A}_\mathrm{R}$. For the customers and rebalancing flows it holds that

$$\sum_{i:(i,j)\in\mathscr{A}} f_m(i,j) + \mathbb{1}_{j=o_m} \cdot \alpha_m = \sum_{k:(j,k)\in\mathscr{A}} f_m(j,k) + \mathbb{1}_{j=d_m} \cdot \alpha_m \qquad \forall m \in \mathscr{M},\ j \in \mathscr{V} \tag{1a}$$

$$\sum_{i:(i,j)\in\mathscr{A}_\mathrm{R}} \left(f_0(i,j) + \sum_{m\in\mathscr{M}} f_m(i,j) \right) = \sum_{k:(j,k)\in\mathscr{A}_\mathrm{R}} \left(f_0(j,k) + \sum_{m\in\mathscr{M}} f_m(j,k) \right) \qquad \forall j \in \mathscr{V}_\mathrm{R} \tag{1b}$$

$$f_m(i,j) \geq 0 \qquad \forall m \in \mathscr{M}, (i,j) \in \mathscr{A} \tag{1c}$$

$$f_0(i,j) \geq 0 \qquad \forall (i,j) \in \mathscr{A}_\mathrm{R}, \tag{1d}$$

where $\mathbb{1}_{j=x}$ is the boolean indicator function. Specifically, we preserve flow conservation for every transportation demand in (1a). Analogously, we guarantee flow conservation for vehicles on every road node with (1b) and ensure non-negativity of flows by (1c) and (1d).

B. Travel Time and Road Congestion

The modeling of travel times and road congestion heavily affects the tractability of our solution approach but also the accuracy of its results. We choose a modeling approach that is a good trade-off between accuracy and tractability for mesoscopic analysis.

We assume a constant walking speed for pedestrian arcs, infer travel times for public transit from the public transit schedules, and use constant average values for mode-switching arcs. On roads, congestion strongly influences travel times. For network flow models it is common practice to scale the nominal travel time at free-flow speed with a volume delay function. In this paper, we use the Bureau of Public Roads (BPR) function [35], which has the form $F_\mathrm{BPR}(x) = 1 + 0.15 \cdot x^4$, with x representing the ratio between the nominal capacity of the road and the vehicles' flow traversing it. With this model, the travel time for the exogenous traffic flow u^R_{ij} in the absence of AMoD vehicles on a road arc $(i,j) \in \mathscr{A}_\mathrm{R}$ results to

$$t_{ij} = t^N_{ij} \cdot F_\mathrm{BPR}\left(u^\mathrm{R}_{ij} / c^\mathrm{R}_{ij}\right), \tag{2}$$

with the travel time at free-flow speed denoted as t^N_{ij} and the nominal road capacity c^R_{ij} depending on the free-flow speed, the number of lanes, and the space occupied by one vehicle.

We model the travel time at free-flow speed on urban road arcs with the following rationale: We assume a car to traverse an arc starting from idling, accelerating with a maximum acceleration a_max until reaching the free-flow speed $v_{\mathrm{max},ij}$, driving with this speed until almost the end of the arc, and finally decelerating with maximum deceleration $-a_\mathrm{max}$, in order to stop at the end of the arc. Capturing acceleration and deceleration in this way allows to account for the driving behaviour also in between crossroads. Consider a road arc with total length s_{ij}. The car reaches the maximum velocity at time $t^*_{ij} = v_{\mathrm{max},ij}/a_\mathrm{max}$ such that, assuming the total travel time to be larger than the total acceleration and deceleration time, the total travel distance on the road arc must satisfy

$$s_{ij} = -\frac{v^2_{\mathrm{max},ij}}{a_\mathrm{max}} + v_{\mathrm{max},ij} \cdot t^N_{ij} \quad \forall (i,j) \in \mathscr{A}_\mathrm{R},$$

yielding the free-flow travel time on road arcs

$$t^N_{ij} = \frac{v_{\mathrm{max},ij}}{a_\mathrm{max}} + \frac{s_{ij}}{v_{\mathrm{max},ij}} \quad \forall (i,j) \in \mathscr{A}_\mathrm{R}. \tag{3}$$

We compare our model with the guidelines provided by the Transportation Research Board (TRB) [36, Ch. 15] in Fig. 2, showing that the travel time computed with our model for different values of link length s_{ij} and free-flow speed $v_{\mathrm{max},ij}$ is in good agreement with the TRB guidelines.

In order to limit the endogenous impact of the AMoD vehicles on road traffic and travel time, we impose a maximum capacity threshold $c^{\mathrm{R,th}}_{ij}$ as

$$f_0(i,j) + \sum_{m\in\mathscr{M}} f_m(i,j) + u^\mathrm{R}_{ij} \leq c^{\mathrm{R,th}}_{ij} \quad \forall (i,j) \in \mathscr{A}_\mathrm{R}. \tag{4}$$

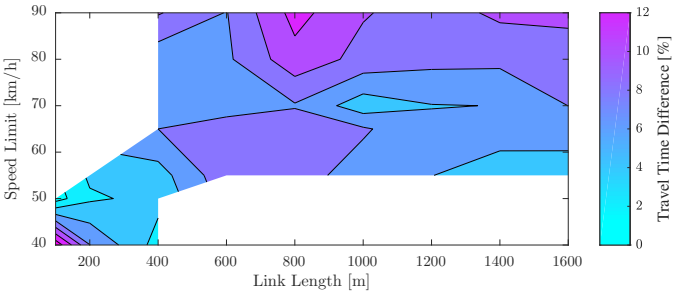

Fig. 2. Relative free-flow travel time difference from the TRB model for road arcs with different length and free-flow speed.

Specifically, we choose the capacity threshold so that the presence of endogenous AMoD traffic does not increase travel time by more than $\Delta r_{\text{time}} \cdot t_{ij}^N$. That is, for each arc $(i,j) \in \mathscr{A}_{\text{R}}$ the capacity threshold $c_{ij}^{\text{R,th}}$ satisfies

$$t_{ij} + \Delta r_{\text{time}} \cdot t_{ij}^N = t_{ij}^N \cdot F_{\text{BPR}}\left(c_{ij}^{\text{R,th}} / c_{ij}^{\text{R}}\right). \tag{5}$$

Choosing a sufficiently small Δr_{time}, the total travel time results from the combination of exogenous and endogenous vehicle flows and can be approximated with its upper bound at full capacity as

$$t_{ij} = t_{ij}^N \cdot F_{\text{BPR}}\left(c_{ij}^{\text{R,th}} / c_{ij}^{\text{R}}\right) \quad \forall (i,j) \in \mathscr{A}_{\text{R}}. \tag{6}$$

C. Energy Consumption of AMoD Vehicles

We compute the energy consumption for AMoD vehicles by applying the general approach from [37] to the New York City urban driving cycle [38]. To account for different mean velocities, we multiply the time and divide the velocity trajectory of the driving cycle by the scaling factor $r_{\text{scale}} = \frac{t_{ij}}{s_{ij}} \cdot \frac{s_{\text{cycle}}}{t_{\text{cycle}}}$, where s_{cycle} and t_{cycle} represent the nominal spatial and temporal length of the driving cycle. Assuming the vehicles to be powered either by gasoline engines with start and stop capabilities or by electrical motors, we compute the energy consumption E in terms of fuel energy E_{f} or electrical battery energy E_{b}, respectively, and assign it to the road arc $(i,j) \in \mathscr{A}_{\text{R}}$ as $e_{ij} = E \cdot s_{ij} / s_{\text{cycle}}$.

For a driving cycle consisting of a velocity, an acceleration and a road inclination trajectory over time $\big(v(t), a(t), \vartheta(t)\big)$, the requested power at the wheels is

$$\begin{aligned} P_{\text{req}}(t) = \Big(& m_{\text{v}} \cdot a(t) + m_{\text{v}} \cdot g \cdot \sin\big(\vartheta(t)\big) + c_{\text{r}} \cdot m_{\text{v}} \cdot g \cdot \cos\big(\vartheta(t)\big) \\ & + \frac{\rho_{\text{air}}}{2} \cdot c_{\text{d}} \cdot A_{\text{f}} \cdot v(t)^2 \Big) \cdot v(t), \end{aligned} \tag{7}$$

where the first term in brackets accounts for the acceleration of the vehicle, the second one is the gravitational force, the third one the rolling friction force, and the last one the aerodynamic drag. Hereby, m_{v} is the mass of the car, g the gravitational acceleration, c_{r} the rolling friction coefficient, ρ_{air} the air density, c_{d} the aerodynamic drag coefficient, and A_{f} the frontal area of the car. Most of these parameters vary depending on the vehicle type. Assuming a constant final drive efficiency η_{fd}, the power provided by the propulsion system P_{p} is

$$P_{\text{p}} = \begin{cases} \frac{1}{\eta_{\text{fd}}} \cdot P_{\text{req}} & \text{if } P_{\text{req}} \geq 0 \\ \eta_{\text{fd}} \cdot P_{\text{req}} + P_{\text{brk}} & \text{if } P_{\text{req}} < 0, \end{cases} \tag{8}$$

where $P_{\text{brk}} \geq 0$ is the braking power exerted by the hydraulic brakes. We model the power provided to the auxiliaries (e.g., heating, ventilation, air-conditioning, ECU, hydraulic brakes, etc.) with a vehicle-size-dependent constant power P_{aux}. Finally, we compute the energy consumption of Internal Combustion Engine Vehicle (ICEV) and Battery Electric Vehicle (BEV) as follows.

1) ICEV: In this case, the car is powered by an internal combustion engine which power needs to match the propulsive power defined in (8), i.e., $P_{\text{e}} = P_{\text{p}}$. We model the engine using a power Willans approximation [39] as $P_{\text{e}} = \eta_{\text{e}} \cdot P_{\text{f}} - P_{\text{e,0}}$, where P_{e} is the engine power, η_{e} the internal efficiency of the engine, P_{f} the fuel power, and $P_{\text{e,0}}$ the engine friction power. Assuming start and stop capabilities, the fuel power is

$$P_{\text{f}} = \begin{cases} 0 & \text{if } P_{\text{req}} \leq 0 \\ \frac{1}{\eta_{\text{e}}} \cdot \left(P_{\text{e}} + P_{\text{e,0}} + \frac{t_{\text{on}}}{t_{\text{tot}}} \cdot P_{\text{aux}}\right) & \text{if } P_{\text{req}} > 0, \end{cases} \tag{9}$$

where t_{tot} is the length of the driving cycle and t_{on} captures the amount of time the ICE is on as $\int_0^{t_{\text{tot}}} \mathbb{1}_{P_{\text{req}}(t)>0} \, \text{d}t$. The fuel energy consumption is then

$$E_{\text{f}} = \int_0^{t_{\text{tot}}} P_{\text{f}}(t) \, \text{d}t. \tag{10}$$

2) BEV: In this case, the car is powered solely by an electrical motor such that the mechanical power needs to match the propulsive power (8) as $P_{\text{m}} = P_{\text{p}}$. We model the efficiency of the electric motor in a piecewise affine manner, to distinguish motor and generator operation. Specifically, the electrical motor power P_{el} is related to the mechanical motor power as

$$P_{\text{el}} = \begin{cases} \frac{1}{\eta_{\text{m}}} \cdot P_{\text{m}} & \text{if } P_{\text{m}} \geq 0 \\ \eta_{\text{g}} \cdot P_{\text{m}} & \text{if } P_{\text{m}} < 0, \end{cases} \tag{11}$$

where η_{m} and η_{g} represent the electrical motor and generator efficiency, respectively. The power delivered by the battery P_{b} is the sum of the electrical motor power and the auxiliary power as $P_{\text{b}} = P_{\text{el}} + P_{\text{aux}}$. We model the internal power drawn from the battery P_{i} in a piecewise affine manner, distinguishing between battery charge and discharge. In particular, we have that

$$P_{\text{i}} = \begin{cases} \frac{1}{\eta_{\text{dis}}} \cdot P_{\text{b}} & \text{if } P_{\text{b}} \geq 0 \\ \eta_{\text{chg}} \cdot P_{\text{b}} & \text{if } P_{\text{b}} < 0, \end{cases} \tag{12}$$

where η_{chg} and η_{dis} represent the battery charging and discharging efficiency, respectively. Finally, the electrical energy consumption is

$$E_{\text{b}} = \int_0^{t_{\text{tot}}} P_{\text{i}}(t) \, \text{d}t. \tag{13}$$

D. I-AMoD Objective and Optimization Problem

Our goal is to maximize the social welfare by minimizing the customers' travel time together with the operational costs incurred by the I-AMoD system. We define commuting costs that depend on the customers' value of time and on operational costs for the AMoD fleet and the public transportation. We assume customers to have the same value of time V_{T}; we define the costs for the AMoD fleet as mileage-dependent costs $V_{\text{D,R}}$ to account for maintenance, depreciation, and the AMoD operators' normal profits (i.e., the profits which compensate

them for their risk and opportunity cost [40]), as well as energy costs V_{E} to account for fuel or electricity consumption. We cumulate all operational costs for the public transportation network as $V_{\mathrm{D,P}}$. Finally, we add a quadratic regularization term with a very small weight V_{Q}. While this term does not appreciably influence the total cost, it does ensure strict convexity for the problem – a key property that enables the design of a socially-optimal pricing and tolling scheme in Section III. The social cost is then

$$\begin{aligned} J_{\mathrm{M}}\big(\{f_m(\cdot,\cdot)\}_m, f_0(\cdot,\cdot)\big) &= V_{\mathrm{T}} \cdot \sum_{m\in\mathscr{M},(i,j)\in\mathscr{A}} t_{ij}\cdot f_m(i,j) \\ &+ \sum_{(i,j)\in\mathscr{A}_{\mathrm{R}}} (V_{\mathrm{D,R}}\cdot s_{ij} + V_{\mathrm{E}}\cdot e_{ij}) \cdot \left(f_0(i,j) + \sum_{m\in\mathscr{M}} f_m(i,j)\right) \\ &+ V_{\mathrm{D,P}} \cdot \sum_{(i,j)\in\mathscr{A}_{\mathrm{P}}} s_{ij} \cdot \sum_{m\in\mathscr{M}} f_m(i,j) \\ &+ V_{\mathrm{Q}} \cdot \left(\sum_{m\in\mathscr{M}} \sum_{(i,j)\in\mathscr{A}} f_m(j,j)^2 + \sum_{(i,j)\in\mathscr{A}_{\mathrm{R}}} f_0(i,j)^2 \right). \end{aligned} \tag{14}$$

We state the I-AMoD optimization problem as follows.

Problem 1 (I-AMoD Optimization Problem)**.** *Given the set of transportation demands $\mathscr{R}$, the optimal customer flows $\{f_m(\cdot,\cdot)\}_m$ and rebalancing flows $f_0(\cdot,\cdot)$ result from*

$$\begin{aligned} \min_{\{f_m(\cdot,\cdot)\}_m, f_0(\cdot,\cdot)} \quad & J_{\mathrm{M}}\big(\{f_m(\cdot,\cdot)\}_m, f_0(\cdot,\cdot)\big) \\ \text{s.t. Eq. (1), Eq. (4), Eq. (14).} \end{aligned} \tag{15}$$

For ease of notation, we reformulate Problem 1 in matrix form. With a slight abuse of notation, let a denote the arc label of an arbitrary arc (i,j). We define

$$\begin{aligned} &[x_{\mathrm{R},m}]_a \coloneqq f_m(i,j), && [x_0]_a \coloneqq f_0(i,j), \\ &[c_{\mathrm{R}}]_a \coloneqq V_{\mathrm{T}}\cdot t_{ij} + V_{\mathrm{D,R}}\cdot s_{ij} + V_{\mathrm{E}}\cdot e_{ij}, && [c_0]_a \coloneqq V_{\mathrm{D,R}}\cdot s_{ij} + V_{\mathrm{E}}\cdot e_{ij}, \\ &[h_{\mathrm{R}}]_a \coloneqq c_{ij}^{\mathrm{R,th}} - u_{ij}^{\mathrm{R}} && \forall (i,j)\in\mathscr{A}_{\mathrm{R}} && \text{(16a)} \\ &[x_{\mathrm{W},m}]_a \coloneqq f_m(i,j), && [c_{\mathrm{W}}]_a \coloneqq V_{\mathrm{T}}\cdot t_{ij} \\ & && \forall (i,j)\in\mathscr{A}_{\mathrm{W}} && \text{(16b)} \\ &[x_{\mathrm{P},m}]_a \coloneqq f_m(i,j), && [c_{\mathrm{P}}]_a \coloneqq V_{\mathrm{T}}\cdot t_{ij} + V_{\mathrm{D,P}}\cdot s_{ij} \\ & && \forall (i,j)\in\mathscr{A}_{\mathrm{P}} && \text{(16c)} \\ &[x_{\mathrm{C},m}]_a \coloneqq f_m(i,j), && [c_{\mathrm{C}}]_a \coloneqq V_{\mathrm{T}}\cdot t_{ij} \\ & && \forall (i,j)\in\mathscr{A}_{\mathrm{C}}, && \text{(16d)} \end{aligned}$$

and denote by $B_{\mathrm{R}} \in \{-1,0,1\}^{|\mathscr{V}_{\mathrm{R}}|\times|\mathscr{A}_{\mathrm{R}}|}$ the incidence matrix of the road graph $\mathscr{G}_{\mathrm{R}}$. Finally, we define $x_m \coloneqq (x_{\mathrm{R},m}, x_{\mathrm{W},m}, x_{\mathrm{P},m}, x_{\mathrm{C},m})$ and denote by $B \in \{-1,0,1\}^{|\mathscr{V}|\times|\mathscr{A}|}$ the incidence matrix of the full graph $\mathscr{G}$. With this notation, we reformulate Problem 1 using the incidence matrix to express flow conservation constraints.

Problem 2 (I-AMoD Optimization Problem Revisited)**.** *Given the set of transportation demands $\mathscr{R}$, the optimal customer flows $\{x_m\}_m$ and rebalancing flows x_0 result from the quadratic optimization problem*

$$\begin{aligned} \min_{\substack{\{x_{\mathrm{R},m}, x_{\mathrm{W},m},\\ x_{\mathrm{P},m}, x_{\mathrm{C},m}\}_m, x_0}} \quad & \sum_{i\in\mathscr{M}} V_{\mathrm{Q}}\cdot x_{\mathrm{R},m}^\top x_{\mathrm{R},m} + c_{\mathrm{R}}^\top x_{\mathrm{R},m} + V_{\mathrm{Q}}\cdot x_{\mathrm{W},m}^\top x_{\mathrm{W},m} + c_{\mathrm{W}}^\top x_{\mathrm{W},m} \\ & + V_{\mathrm{Q}}\cdot x_{\mathrm{P},m}^\top x_{\mathrm{P},m} + c_{\mathrm{P}}^\top x_{\mathrm{P},m} + V_{\mathrm{Q}}\cdot x_{\mathrm{C},m}^\top x_{\mathrm{C},m} + c_{\mathrm{C}}^\top x_{\mathrm{C},m} \\ & + V_{\mathrm{Q}}\cdot x_0^\top x_0 + c_0^\top x_0 && && \text{(17a)} \\ \text{s.t. } & Bx_m = b_m && \forall m\in\mathscr{M} && \text{(17b)} \\ & B_{\mathrm{R}}\left(\sum_{m\in\mathscr{M}} x_{\mathrm{R},m} + x_0\right) = 0 && && \text{(17c)} \\ & \sum_{m\in\mathscr{M}} x_{\mathrm{R},m} + x_0 \le h_{\mathrm{R}} && && \text{(17d)} \\ & x_m \ge 0,\ x_0 \ge 0 && \forall m\in\mathscr{M}, && \text{(17e)} \end{aligned}$$

where $[b_m]_i = -\alpha_m$ for the origin node, $[b_m]_i = +\alpha_m$ for the destination node, and $[b_m]_i = 0$ otherwise. The linear terms in (17a) *account for the cost incurred when travelling on the corresponding digraph by customers and rebalancing vehicles. The quadratic terms act as regularizers. Conservation of customers and vehicles is guaranteed by* (17b) *and* (17c) *respectively. Constraint* (17d) *captures road congestion.*

Lemma II.2. *Problem 1 and Problem 2 are solution equivalent, i.e., the optimal solution of Problem 1 denotes the optimal solution of Problem 2 and vice versa.*

Proof. Substituting cost vectors and variables in (17a) directly shows the equivalence. With

$$[Bx_m]_i = \sum_{i:(i,j)\in\mathscr{A}} f_m(i,j) - \sum_{j:(i,j)\in\mathscr{A}} f_m(i,j) = \begin{cases} +\alpha_m & \text{if } i = d_m \\ -\alpha_m & \text{if } i = o_m \\ 0 & \text{else,} \end{cases}$$

(1a) is equivalent to (17b) and (17c)–(17e) follow analogously. □

Lemma II.3. *Problems 1 is feasible and has a unique solution.*

Proof. Feasibility: The proof is constructive. By Assumption 2, requests originate and end on nodes in the walking graph. By Assumption 1 the walking graph is strongly connected, therefore, there always exists a path connecting any pair of nodes $o,d\in\mathscr{V}_{\mathrm{W}}$. Also, the capacity of all arcs in $\mathscr{A}_{\mathrm{W}}$ is infinite. For each customer request $r=(o,d,\alpha)$ wishing to travel from node o to node d, select a path p connecting o to d containing only arcs in the pedestrian graph, and set flows flows $f_m(\cdot,\cdot)$ equal to α for all arcs in the path and zero otherwise. The resulting flow is a feasible solution. *Uniqueness*: Solution uniqueness follows directly from strict convexity, the affinity of the constraints of Problem 2, and Lemma II.2. □

E. Discussion

A few comments are in order. First, we consider a time-invariant transportation demand. This assumption is in order if the requests change slowly compared to the average travel time of individual trips, as is often observed in densely populated urban environments [41]. Second, we do not explicitly account for the stochastic nature of exogenous traffic and of the customer arrival process. Given the mesoscopic perspective of our study, this deterministic representation is in order as it captures such stochastic processes *on average* [9]. Third, we

allow fractional customer and vehicle flows. We show in Section IV-B1, that the resulting accuracy loss is negligible for the mesoscopic perspective of our study. Moreover, Problem 2 can be solved in polynomial time with off-the-shelf optimization algorithms providing global optimality guarantees for the solution found. For real-time applications, randomized sampling methods can be used to compute integer-valued flows from fractional flows, yielding near-optimal routes for individual vehicles and customers [42, Ch. 4], whilst new information can be accounted for as it is revealed through a receding-horizon framework. Fourth, we adopt a threshold model to capture road congestion. On the one hand, such a model allows us to *contain* the impact of AMoD vehicles on road traffic, whilst on the other hand it can be expressed as a linear inequality constraint. Methods to account for the impact of endogenous traffic on travel time via volume-delay functions such as the BPR function [35] have been presented in [17] for the AMoD-only problem, whereas convex approximations and relaxations readily applicable to the I-AMoD case can be found in [16]. In this paper, we focus on scenarios where the AMoD traffic cannot exceedingly impact the road network in order to devise road tolling schemes that provide these conditions. Choosing a sufficiently small Δr_{time} allows us to use the upper bound of (6) as an estimate travel time for road arcs. Fifth, we assume exogenous traffic not to be affected by the endogenous AMoD vehicle routes. This assumption is also acceptable for small values of Δr_{time}, capturing the fact that vehicles follow similar routes under similar traffic conditions. We leave the game-theoretical extension accounting for reactive exogenous traffic to future research work. Sixth, we allow AMoD vehicles to transport one customer at a time. Such an assumption is in line with current trends in mobility-on-demand systems, such as taxis, Lyft and Uber. The extension to ride-sharing AMoD requires the adoption of integer-valued flows and time-expansions of the whole transportation network, resulting in prohibitory trade-offs between computational times and model accuracy [43]. Finally, for the sake of simplicity, we consider customers to have identical preferences in terms of value of time and travel comfort. However, the model proposed in this paper can be readily extended to capture distinct classes of customers, each characterized by a different network flow associated with specific preferences.

III. A Pricing and Tolling Scheme for I-AMoD

In Section II, we assume that the objectives of all stakeholders are aligned with the global objective of maximizing social welfare. In reality, stakeholders are selfish, i.e., customers maximize their private welfare, whilst AMoD fleet operators maximize their profits. In this section, we propose a road tolling scheme to align the goals of self-interested agents with the objective of maximizing social welfare (cf. Section II-D). Section III-A formally introduces the self-interested agents participating in the I-AMoD market, while Section III-B details our road tolling scheme and Section III-C proves its alignment with the social optimum.

A. Self-interested Agents

We model the I-AMoD market as a perfect market with three types of agents: The municipal transportation authority, I-AMoD customers, and AMoD operators. Assuming a perfect I-AMoD market, neither individual customers nor AMoD operators can unilaterally influence the transportation prices which result from the market equilibrium [40].

The municipal transportation authority sets fares in the subway system and road tolls in the road network aiming at maximal social welfare. Prices in the public transportation network are set to cover the operational cost of the transportation system, whereas road tolls can be interpreted as congestion surcharges. Specifically, the transportation authority sets a fare $p_{\text{P}}(i,j)$ for each arc $(i,j) \in \mathscr{A}_{\text{P}}$ in the public transportation network and a toll $\tau_{\text{R}}(i,j)$ for each arc $(i,j) \in \mathscr{A}_{\text{R}}$ in the road network.

I-AMoD customers serve their mobility requests $m \in \mathscr{M}$ by selecting an intermodal route from their origins to their destinations. From our mesoscopic perspective, route selection consists of choosing a commodity flow $f_m(\cdot,\cdot)$ satisfying continuity as in Eq. (1a). We neglect common user-centric modeling approaches that account for individual cost functions: In line with current practice [44], we assume that customers select their routes by using navigation apps which compute routes by considering an aggregate model of the customers' preferences. Specifically, we set a customer's objective as the maximization of her welfare, defined as the sum of the travel time multiplied by the value of time V_{T} and the cost of her trip as the cumulative sum of the fares paid along the route: $p_{\text{P}}(i,j)$ for each traversed arc $(i,j) \in \mathscr{A}_{\text{P}}$ in the public transportation network, as set by the municipal authority, and $p_{\text{R}}(i,j)$ for each road arc $(i,j) \in \mathscr{A}_{\text{R}}$ traveled with an AMoD vehicle. Considering a negligible quadratic regularization term (cf. Eq. 14) for the sake of consistency, a customer's I-AMoD navigation app solves the following problem.

Problem 3 (I-AMoD Customers Optimization Problem)**.** *Given a transportation request $r_m = (o_m, d_m, \alpha_m) \in \mathscr{R}$, I-AMoD customers' routes result from*

$$\begin{aligned}\min_{f_m(\cdot,\cdot)} \; & V_{\text{T}} \cdot \sum_{(i,j)\in\mathscr{A}} t_{ij} \cdot f_m(i,j) + V_{\text{Q}} \cdot \sum_{(i,j)\in\mathscr{A}} f_m(j,j)^2 \\ & + \sum_{(i,j)\in\mathscr{A}_{\text{R}}} p_{\text{R}}(i,j) \cdot f_m(i,j) + \sum_{(i,j)\in\mathscr{A}_{\text{P}}} p_{\text{P}}(i,j) \cdot f_m(i,j) \\ \text{s.t. } & \text{Eq. (1a), Eq. (1c),}\end{aligned} \tag{18}$$

The first term in the cost function corresponds to the customer's value of time, the third term denotes the arc-based charge in the network, and the fourth term is the fare paid to the subway network.

AMoD operators service customers and control the rebalancing vehicles' routes to ensure that vehicles are available to service customer requests. Without loss of generality, we fold the AMoD operators into a unique operator, paying tolls $\tau_{\text{R}}(i,j)$ to the municipal authority and levying fares $p_{\text{R}}(i,j)$ from the customers for each road arc traversed. As the AMoD operator is unable to influence the AMoD prices $p_{\text{R}}(i,j)$ in a perfect market, the goal of maximizing revenue is equivalent to

the goal of minimizing operating expenses. Again, we include a negligible quadratic regularization term to ensure strict convexity such that the AMoD operator solves the following problem.

Problem 4 (AMoD Operator Optimization Problem)**.** *Given the customer flows $\{f_m(\cdot,\cdot)\}_m$, the optimal AMoD rebalancing flows $f_0(\cdot,\cdot)$ result from*

$$\min_{f_0(\cdot,\cdot)} \sum_{(i,j)\in\mathscr{A}_\mathrm{R}} (V_{\mathrm{D,R}}\cdot s_{ij} + V_\mathrm{E}\cdot e_{ij} + \tau_\mathrm{R}(i,j))\cdot f_0(i,j) + V_\mathrm{Q}\cdot f_0(i,j)^2$$
$$\text{s.t. Eq. (1b), Eq. (1c).} \tag{19}$$

Analogous to Section II-D, we reformulate Problems 3 and 4 in matrix notation. Let $[\tilde{c}_\mathrm{R}]_a := V_\mathrm{T}\cdot t_{ij} + p_\mathrm{R}(i,j)$ and $[\bar{c}_0]_a := V_{\mathrm{D,R}}\cdot s_{ij} + V_\mathrm{E}\cdot e_{ij} + \tau_\mathrm{R}(i,j)\,\forall (i,j)\in\mathscr{A}_\mathrm{R}$, $[\tilde{c}_\mathrm{W}]_a := V_\mathrm{T}\cdot t_{ij}\,\forall(i,j)\in\mathscr{A}_\mathrm{W}$, $[\tilde{c}_\mathrm{P}]_a := V_\mathrm{T}\cdot t_{ij}\,\forall(i,j)\in\mathscr{A}_\mathrm{P}$, and $[\tilde{c}_\mathrm{C}]_a := V_\mathrm{T}\cdot t_{ij}\,\forall(i,j)\in\mathscr{A}_\mathrm{C}$.

Problem 5 (I-AMoD Customers Optimization Problem Revisited)**.** *Given a transportation request $r_m = (o_m, d_m, \alpha_m)\in\mathscr{R}$, I-AMoD a customer's route results from*

$$\min_{\substack{x_{\mathrm{R},m},x_{\mathrm{W},m},\\ x_{\mathrm{P},m},x_{\mathrm{C},m}}} V_\mathrm{Q}\cdot x_{\mathrm{R},m}^\top x_{\mathrm{R},m} + \tilde{c}_\mathrm{R}^\top x_{\mathrm{R},m} + V_\mathrm{Q}\cdot x_{\mathrm{W},m}^\top x_{\mathrm{W},m} + \tilde{c}_\mathrm{W}^\top x_{\mathrm{W},m} \tag{20a}$$
$$+ V_\mathrm{Q}\cdot x_{\mathrm{P},m}^\top x_{\mathrm{P},m} + \tilde{c}_\mathrm{P}^\top x_{\mathrm{P},m} + V_\mathrm{Q}\cdot x_{\mathrm{C},m}^\top x_{\mathrm{C},m} + \tilde{c}_\mathrm{C}^\top x_{\mathrm{C},m}$$
$$\text{s.t. } Bx_m = b_m \tag{20b}$$
$$x_m \geq 0, \tag{20c}$$

where the linear terms in (20a) *capture travelling cost, the quadratic terms comprise the regularization, and the matrix equality constraint* (20b) *ensures flow conservation.*

Problem 6 (AMoD Operator Optimization Problem Revisited)**.** *Given the customer flows $\{x_m\}_m$, the optimal AMoD rebalancing flows result from*

$$\min_{x_0} V_\mathrm{Q}\cdot x_0^\top x_0 + \bar{c}_0^\top x_0 \tag{21a}$$
$$\text{s.t. } B_\mathrm{R}\left(\sum_{m\in\mathscr{M}} x_{\mathrm{R},m} + x_0\right) = 0 \tag{21b}$$
$$x_0 \geq 0. \tag{21c}$$

While the linear term in the (21a) *represents travelling cost, the quadratic term denotes the regularization, and* (21b) *enforces vehicle conservation.*

Lemma III.1. *Problem 3 and Problem 5 are equivalent.*

Lemma III.2. *Problem 4 and Problem 6 are equivalent.*

The proofs of Lemmas 5 and 6 are identical to the proofs of Lemma II.2.

Lemma III.3. *Problem 3 is feasible and has a unique solution.*

Proof. Let $\left\{\{f_m(\cdot,\cdot)\}_m^\star, f_0(\cdot,\cdot)^\star\right\}$ be a guaranteed feasible solution to Problem 1 (cf. Lemma II.3). Then, $f_m(\cdot,\cdot)^\star$ is a feasible, yet suboptimal, solution to Problem 3. Hence, Problem 3 is feasible. Uniqueness follows analogously to Lemma II.3 . □

Lemma III.4. *Problem 4 is feasible and has a unique solution.*

Proof. Since there are no capacity constraints, the proof follows analogously to Lemma II.3 and III.3. □

B. A Pricing and Tolling Scheme

The subway fares $p_\mathrm{P}(i,j)$ and the road tolls $\tau_\mathrm{R}(i,j)$ are control variables that the welfare-minded municipal authority can adjust to steer self-interested customers and the AMoD operator towards maximizing social welfare as defined in Section II-D.

We denote the dual multipliers associated with the road capacity constraint (17d) as μ_cR, and the dual multipliers associated with the vehicle balance constraints (17b) as λ_R. To ease the notation we let $\mu_\mathrm{cR}(i,j) := [\mu_\mathrm{cR}]_a$ and $\lambda_\mathrm{R}(i) := [\lambda_\mathrm{R}]_i$.

We propose the following pricing and tolling scheme: The subway fares are set equal to the the public transit operational cost as

$$p_\mathrm{P}(i,j) = V_{\mathrm{D,P}}\cdot s_{ij}, \tag{22}$$

whilst the road tolls are chosen equal to the road congestion multipliers as

$$\tau_\mathrm{R}(i,j) = \mu_\mathrm{cR}(i,j). \tag{23}$$

C. A General Equilibrium

Given the market presented in Section III-A, we can define its general economic equilibrium as follows.

Definition III.5 (General Economic Equilibrium)**.** *A solution $\left\{\{f_m(\cdot,\cdot)\}_m^\star, f_0(\cdot,\cdot)^\star\right\}$ and a set of prices $\{p_\mathrm{P}(i,j), \tau_\mathrm{R}(i,j), p_\mathrm{R}(i,j)\}$ form a general economic equilibrium if and only if 1) $f_m(\cdot,\cdot)^\star$ is a solution to Problem 3 for all $m\in\mathscr{M}$, 2) $f_0(\cdot,\cdot)^\star$ is a solution to Problem 4, and 3) the economic profit of each AMoD operator is zero (that is, the operator's revenue equals its costs).*

Remark III.6. *The requirement that, at equilibrium, the economic profit of each AMoD operator equals the costs is characteristic of perfect markets, where sellers have no economic surplus [40]. The condition does not imply that AMoD operators receive no profit whatsoever; indeed, the cost $V_{\mathrm{D,R}}$ captures the operators' normal profits, which compensate them for their risk and opportunity cost. This condition has an intuitive interpretation. At equilibrium, no operator can lower its prices, since it would then be better off leaving the market; at the same time, no operator can increase its prices, since all customers would prefer to be served by a cheaper competitor.*

Assume the AMoD road prices $p_\mathrm{R}(i,j)$ equal to the sum of the vehicles' operating costs, the road tolls, and the origin and destination prices as

$$p_\mathrm{R}(i,j) = V_{\mathrm{D,R}}\cdot s_{ij} + V_\mathrm{E}\cdot e_{ij} + \tau_\mathrm{R}(i,j) + \lambda_\mathrm{R}(i) - \lambda_\mathrm{R}(j). \tag{24}$$

Remark III.7. *The AMoD origin-destination prices $\sum_{(i,j)\in\mathscr{A}_\mathrm{R}}(\lambda_\mathrm{R}(i) - \lambda_\mathrm{R}(j))f_m(i,j)$ for a request $r = (o,d,\alpha)$ simplify to $\lambda_\mathrm{R}(o) - \lambda_\mathrm{R}(d)$, capturing the marginal cost incurred to rebalance the system due to the request.*

The following theorem shows that the pricing and tolling scheme proposed in Section III-B ensures that an optimal solution to the I-AMoD Problem 1 coincides with a general economic equilibrium for the market.

Theorem III.8 (Optimal Pricing and Tolling Scheme)**.** *Consider the optimal solution $\left\{\{f_m(\cdot,\cdot)\}_m^\star, f_0(\cdot,\cdot)^\star\right\}$ to the*

I-AMoD problem. Also, consider a perfect market where self-interested customers plan their routes with a navigation app solving Problem 3, a self-interested AMoD operator plans rebalancing routes by solving Problem 4, and the municipal transportation authority sets public transit prices and road tolls according to (22)–(23)*. Then, the optimal solution* $\left\{\{f_m(\cdot,\cdot)\}_m^\star, f_0(\cdot,\cdot)^\star\right\}$ *and the prices* (24) *are a general economic equilibrium for the I-AMoD market; that is:*

1) $f_m(\cdot,\cdot)^\star$ *is an optimal solution to Problem 3;*
2) $f_0(\cdot,\cdot)^\star$ *is an optimal solution to Problem 4;*
3) *the AMoD operators's revenue equals its costs (up to the regularization term).*

Proof Sketch. The proof relies on showing that satisfaction of the KKT conditions for the I-AMoD Problem 2 implies satisfaction of the KKT conditions for the customers' optimal routing Problem 5 and the KKT conditions for the AMoD operator's optimal rebalancing Problem 6. We provide a rigorous proof in Appendix B. □

D. Discussion

A few comments are in order. First, in the setting of a general equilibrium, we assume that the AMoD operators have no pricing power, i.e., no individual AMoD operator is able to single-handedly influence the customers' fares. This assumption holds if multiple operators of similar size compete for customers' transportation demands, and is arguably realistic in several urban environments. For reference, no fewer than five app-based mobility-on-demand operators (Uber, Lyft, Juno, Curb and Arro) currently offer mobility-on-demand services in Manhattan. Second, in this paper, the operations of all AMoD operators are captured through a single rebalancing flow and a single set of customer-carrying flows on road arcs for simplicity and ease of notation. However, the model does *not* assume that a single AMoD operator is present. Indeed, a treatment where different operators control different subsets of vehicles, each associated with a rebalancing flow, would result in the same equilibrium. Note that customers, not AMoD operators, choose the operator by selecting the customer-carrying flows $\{f_m(\cdot,\cdot)\}_m^\star$, and the operators do not compete on prices. Therefore, Theorem III.8 still holds as the operation of single AMoD operators are not coupled, and Problem 4 can be decomposed into subproblems, one for each operator. Third, we assume that the routes followed by customer-carrying AMoD vehicles are set by the customers themselves through the navigation apps. In practical implementations, the customers may be able to choose only among a limited set of possible routes, for example between a direct route that incurs congestion tolls and a longer, less congested and thus cheaper route. Such more sophisticated route selection models are left for future research. Fourth, we use the cost function (18) to model customers' behaviour. Although such an approach does not entail the level of detail of a user-centric approach [30, Ch. 4], it suffices for the mesoscopic perspective of this study. Finally, Theorem III.8 shows that the socially optimal solution can be achieved as a general economic equilibrium. However, it does not prove that the flows $\left\{\{f_m(\cdot,\cdot)\}_m^\star, f_0(\cdot,\cdot)^\star\right\}$ and the prices $\{p_{\mathrm{P}}(i,j), \tau_{\mathrm{R}}(i,j), p_{\mathrm{R}}(i,j)\}$ are the *only* equilibrium for the system; indeed, other equilibria may exist that result in higher societal cost compared to the socially optimal solution. To overcome this, we envision that the market could be steered towards the socially optimal equilibrium through the introduction of a non-profit *market-making* entity whose role is to match supply with demand while steering the system towards the socially-optimal equilibrium, akin to the role of non-profit Independent System Operators in electricity markets.

TABLE I
REQUESTS IN BERLIN AND NYC.

	NYC	Berlin
M	8,658	2,646
$\sum_{m\in\mathscr{M}} \alpha_m$	44.943 1/s	3.771 1/s
$\sum_{m\in\mathscr{M}} \alpha_m \lVert o_m - d_m \rVert_2 / \sum_{m\in\mathscr{M}} \alpha_m$	2.4 km	4.0 km

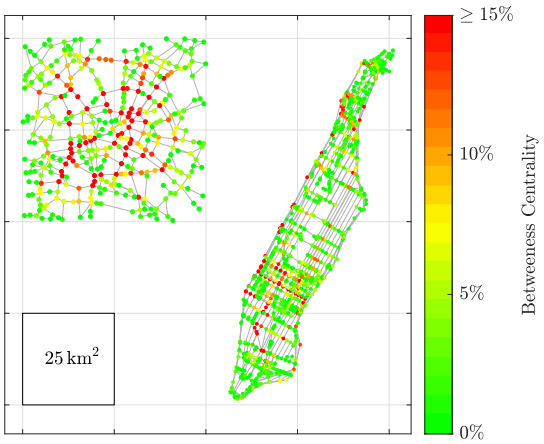

Fig. 3. Measure of the "betweenness" centrality for each node in the road digraphs of Berlin (left) and Manhattan (right). The broad and well-connected structure of Berlin is reflected in several nodes with a high degree of centrality, whereas the elongated shape of Manhattan results in less central nodes.

IV. RESULTS

In this section, we assess the benefits of an I-AMoD system in terms of travel time, costs, and emissions for real-world case studies of NYC and Berlin. Section IV-A details these case studies before we present the optimal solution for the I-AMoD system in Section IV-B. Additionally, we study the change in global cost for different vehicle architectures. Finally, we compare the optimal solution for the I-AMoD system and the AMoD system operating in isolation in Section IV-C.

A. Case Study

We focus on two distinct metropolitan areas, namely, the Manhattan peninsula in NYC, NY, and the city center of Berlin, Germany, as they are complementary in the following parameters: First, the cities differ in the spatial structure of its road system. Fig. 3 shows the "betweenness" centrality of road nodes computed as the probability that a shortest path between two random nodes will traverse the given node. The center of Berlin covers a broad region with several nodes that show a high centrality, whereas Manhattan has less homogeneously distributed central nodes and a more elongated and thinner

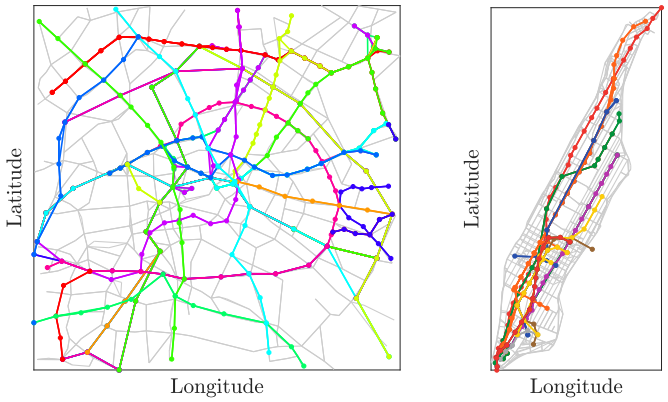

Fig. 4. The public transit network of Berlin (left) and NYC (right).

urban shape. Second, although their surface is nearly equal in size, the amount of trip requests in Manhattan exceeds the amount of trip requests in Berlin by one order of magnitude (cf. Table I). Finally, the geodesic distance between origin-destination-pairs in Berlin is on average twice as long as in NYC.

We derive transportation requests as follows: For Manhattan, we consider the actual 53,932 taxi rides which took place on March 1, 2012 between 6 PM and 8 PM (courtesy of the New York Taxi and Limousine Commission). Although this number of trips is quite large, it represents only a fraction of the travel demand: In 2017 the number of ride-sharing vehicles used in this time period outnumbered yellow cabs by a factor of 5 [45]. Hence, we increase the number of requests by a factor of six to emulate the total demand for ride-hailing services during this time window, obtaining a total of 8,658 origin-destination pairs. For Berlin, we use data from the MATSim [46] Berlin case study [47]. In order to provide a fair comparison, we scale the demand in order to match the demand of Manhattan in terms of requests pro capite. For both cities we derive the road network from OpenStreetMap data [48], and define the capacity of each street to be proportional to the number of lanes multiplied by the road's speed limit [13]. To account for exogenous vehicles on the road, we perform a parametric study, varying the scaling factor of the exogenous road usage u^{R} between 50% and 200% of the nominal road capacity c^{R}. We set the maximum increase in travel time caused by AMoD vehicles Δr_{time} to 5%. We assume the subway network to be the only public transportation system in Manhattan, in line with the fact that the subway network is the dominant public transit mode of the city, whereas in Berlin we also include the S-Bahn and the tram lines (Fig. 4). This way, we provide a first order assessment of I-AMoD. We construct the public transportation digraph using the geographical location of the lines and the stops found in the NYC Open Data database [49] as well as the time schedules of the MTA [50], whereas for Berlin we directly use GTFS data [51]. We set the time to transfer from a road node or a subway stop to a walking node, which models the time required to exit an AMoD vehicle or a subway station, to one minute. We assume that 90 seconds are required to go from a pedestrian to a road node and get into an AMoD vehicle, which is in line with the average time to hail a ride in Manhattan [52]. The time to transfer from a walking node to a subway line equals one minute plus one half of the frequency of the line. We directly relate the energy consumption to the CO_2 emissions based on the current electricity sources of the state of New York [53] and Germany [54], whereas we consider molar mass ratio between CO_2 and CH_2 for gasoline [39]. We compute the energy consumption as in Section II-C for the different vehicles studied, namely, a Lightweight (LW) and a Sport Utility (SU) vehicle. Table II summarizes the remaining parameters used in our case studies and their bibliographic sources.

TABLE II
NUMERICAL DATA FOR THE CASE STUDIES

Parameter	Variable	Value		Source
Maximum time increase	Δr_{time}	5 %		
Value of time	V_{T}	24.40 USD/h		[55]
Vehicle operational cost	$V_{\mathrm{D,R}}$	0.48 USD/mile		[56]
		NYC	Berlin	
Subway operational cost	$V_{\mathrm{D,P}}$	0.47 USD/mile	0.30 USD/mile	[57], [58]
Cost of electricity	V_{E}	0.25 USD/kWh	0.33 USD/kWh	[59], [60]
Cost of gasoline	V_{E}	0.07 USD/kWh	0.18 USD/kWh	
Air density	ρ_{air}	1.25 kg/m^3		
Final drive efficiency	η_{fd}	98 %		[37]
Motor efficiency	$\eta_{\mathrm{m/g}}$	90 %		[37]
Battery efficiency	$\eta_{\mathrm{dis/chg}}$	90 %		[37]
Engine efficiency	η_{e}	40 %		[37]
		LWV	SUV	
Engine drag	$P_{\mathrm{e,0}}$	1.6 kW	4 kW	[37]
Mass of the vehicle	m_{v}	750 kg	2000 kg	[37]
Rolling friction coefficient	c_{r}	0.008	0.017	[37]
Frontal drag coefficient	$c_{\mathrm{d}} \cdot A_{\mathrm{f}}$	0.4 m^2	1.2 m^2	[37]

For each of the scenarios presented in the next Sections IV-B and IV-C, the quadratic optimization Problem 1 was solved on commodity hardware (Intel Core i7, 16 GB RAM) using Gurobi 8.1 in less than 5 minutes.

B. Optimal Solution for the I-AMoD System

In this section, we study the performance achievable by the I-AMoD system by solving Problem 1 considering an AMoD fleet of LW BEVs for different levels of exogenous road usage. Specifically, a 50% baseline road usage corresponds to the case with little traffic, whereas a 200% usage corresponds to the heavy traffic congestion that several urban areas are suffering nowadays. Fig. 5 and 6 show the distance-based modal share in Berlin and NYC for different levels of road usage, together with the average travel time, CO_2 emissions, and monetary cost (J_{M} without regularization terms), scaled with the average distance between origin-destination-pairs. As exogenous road traffic increases, subway utilization grows. With even more traffic the walking distance also increases because the subway cannot fully replace a point-to-point means of transportation. Nevertheless, the travel time and monetary costs remain well contained, while emissions drop significantly. The AMoD modal share follows a similar trend in both cities, whereas the public transit modal share is significantly higher in Berlin, especially under high levels of traffic, as it results from the better and cheaper public transportation system of the European city.

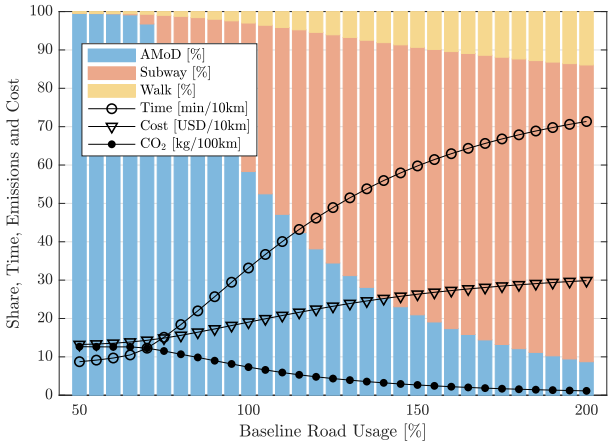

Fig. 5. Distance-based modal share for the I-AMoD system in Berlin with a fleet of lightweight electric vehicles.

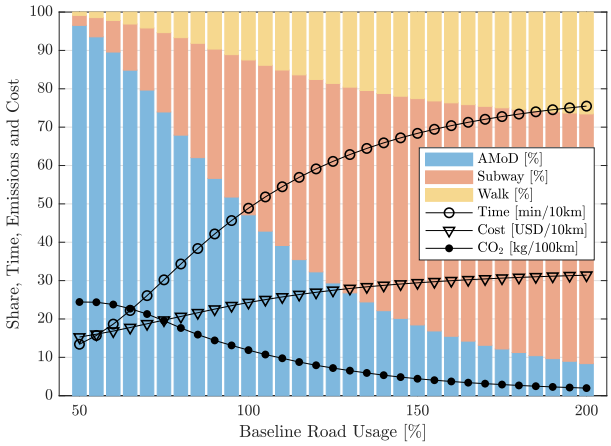

Fig. 6. Distance-based modal share for the I-AMoD system in NYC with a fleet of lightweight electric vehicles.

Overall, Berlin achieves a slightly better performance whilst achieving significantly lower CO_2 emissions due to the higher cost of energy in Europe.

Fig. 7 shows the road tolls $\tau_R(i,j)$ from Eq. (23) for a baseline road usage of 150%. In NYC, the average cost of fares paid due to tolls along trips corresponds to almost 1 USD, which is interestingly in line with the proposal to tax mobility-on-demand vehicles with a congestion surcharge of 2 to 5 USD per trip [33]. In Berlin, the average surcharge would result to 2.45 USD per trip. Considering that in the German city the average distance between origin destination pairs is 1.7 times larger than in NYC, the relative surcharge in Berlin would be about 20% lower.

1) Comparison between Fractional and Integer Solution: To justify our discussion on fractional flows, we study the difference between the fractional and the integer-valued solution by solving Problem 1 imposing additional integer constraints

$$f_m(i,j) \in \mathbb{Z}\ \forall (i,j) \in \mathscr{A}, m \in \mathscr{M}, \quad f_0(i,j) \in \mathbb{Z}\ \forall (i,j) \in \mathscr{A}_R. \quad (25)$$

To improve the model's computational tractability, we set the small quadratic regularization cost V_Q to zero for this comparison, and solve Problem 1 subject to Eq. (25). Table III

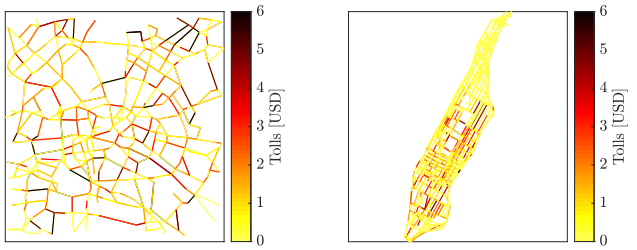

Fig. 7. Optimal road tolls in USD for Berlin (left) and NYC (right).

TABLE III
RELATIVE DIFFERENCE BETWEEN INTEGER AND FRACTIONAL FLOWS.

	Cost	CO_2	Time	Flows
Relative difference [%]	0.0003	0.0015	0.0006	6

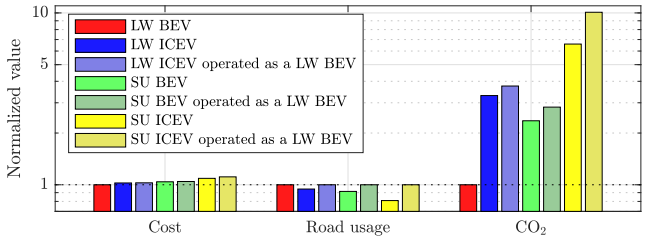

Fig. 8. The optimal solution for different vehicles' configurations for NYC with a 100% baseline road usage shows the impact of the vehicle category and the powertrain type on the optimal system performance and operation.

shows that the resulting integer-valued and fractional solutions coincide in terms of cost but not always in terms of routes, since $V_Q = 0$ results in a non-strictly convex problem with possibly multiple solutions. The negligible difference in terms of costs validates our fractional flow relaxation for mesoscopic studies.

2) Impact of the AMoD Vehicles' Characteristics: We study the impact of different vehicles' categories and powertrain types on the solution. In particular, we consider a BEV and an ICEV from the LW and the SU class, and solve Problem 1 for NYC with a 100% baseline road usage. Fig. 8 shows the results in terms of social cost, road usage, and CO_2 emissions, scaled with respect to the LW BEV case. Moreover, it shows the costs resulting when applying the operational strategies for the LW BEV fleet to fleets with different vehicles. The societal cost increases the more inefficient the vehicle gets, whilst the relative road usage drops. This reveals that a SU BEV is socially worse than a LW ICEV despite having a significantly more efficient powertrain, as trucks are heavier and suffer more drag, resulting in a lower optimal road usage. In fact, the CO_2 emissions caused by the electric SUV are more than twice as large compared to the LW BEV and close to the emissions of the LW ICEV. Additionally, it can be seen that the vehicle type affects the optimal operation of the fleet: Operating the fleets using the solution for LW BEVs is suboptimal in terms of both societal costs and emissions, revealing the importance of considering operational implications and dependencies while taking strategic fleet design decisions and vice versa. Finally, the results for the fleet of SU ICEVs entail emissions one

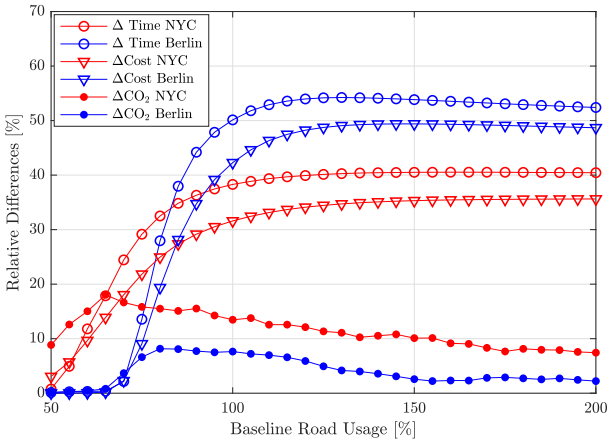

Fig. 9. Relative difference in travel time, cost, emissions, and number of cars between AMoD and I-AMoD in Berlin and in NYC.

order of magnitude larger and slightly higher societal costs, highlighting the inefficiency of such vehicles.

C. Comparison with the Pure AMoD System

We repeat the studies from Section IV-B for the AMoD system operating in isolation to quantify the improvement potential of the intermodal system. We obtain the AMoD solution by simply setting the subway capacity to zero. The "infinite" capacity of the pedestrian network guarantees feasibility even under extremely congested conditions. For the pure AMoD system, the AMoD modal share (namely, the average distance covered by car per trip) observed in both cities is very similar to the one displayed in Fig. 5 and 6. Fig. 9 shows the relative decrease in average travel time, CO_2 emissions, and monetary cost achievable by coordinating the AMoD fleet with the public transportation network. In NYC, as the road availability decreases, the difference in travel time and cost increases up to 40%, while emissions are reduced by almost 20%. In Berlin, the increase in terms of travel time and social cost is 10% larger, whilst CO_2 emissions are less significantly affected. While the road system in Berlin allows for fewer possible routes than the dense and aligned road structure of NYC, the German public transportation system provides a better service, resulting in up to 50% improvements. In both cities, the number of cars decreases by up to 20%.

Fig. 10 shows that due to the severe congestion constraints, the difference in road usage is minor. However, the higher road tolls would cause an average surcharge of almost 2 USD per trip in NYC (2× the I-AMoD surcharge) and 6 USD in Berlin (3× the I-AMoD surcharge). The relatively larger increase in Berlin highlights the lower road density of the city. Overall, I-AMoD results in shorter travel times, fewer vehicles, as well as much lower emissions and tolls.

V. Conclusion

In this paper, we explored the possibility of coordinating different modes of transportation in congested urban environments in order to satisfy travel requests whilst maximizing social welfare. We presented a network flow model for

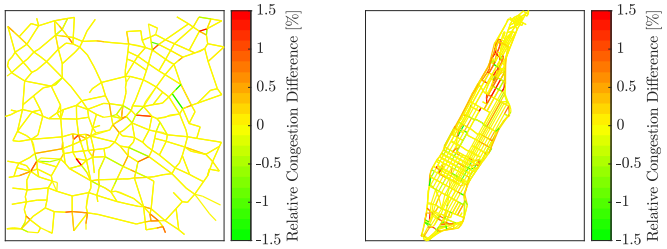

Fig. 10. Relative congestion difference in % for Berlin (left) and NYC (right).

an Autonomous Mobility-on-Demand (AMoD) system that cooperates with the public transportation network. Besides computing optimal customer and vehicle routes, we designed a pricing and tolling scheme that allows to steer selfish agents to the social optimum under the assumption of a perfect market. Through real-world case studies for Manhattan and Berlin, we investigated how the structural differences of their urban transportation networks impact the optimal performance and operation of a centralized intermodal AMoD system. Specifically, our results for Berlin show significantly lower emissions and a slightly better service quality. Berlin's better-connected and cheaper public transportation helps overcome the disadvantages caused by longer average customer trips and, compared to New York, a less dense road structure. Moreover, we characterized the impact of the fleet characteristics by comparing lightweight vehicles and SUVs powered by an electric motor or a gasoline engine. Our numerical studies revealed the vehicle size to have a comparable importance as the powertrain type, whilst underlining the benefit of accounting for the fleet characteristics when devising operational strategies and vice versa. Finally, we showed that optimizing the customers and the rebalancing routes by jointly considering the AMoD fleet and public transit can significantly reduce travel time, costs, number of vehicles and emissions. The optimal road tolls computed within our framework are quantitatively in line with the surcharges for ride-hailing trips recently discussed by the New York City municipal authority and significantly lower than for the AMoD system in isolation.

This work opens the field for several research directions. First, we would like to design an operational algorithm to compute the optimal customer and rebalancing routes at the microscopic level in real-time. Second, it is of interest to extend our model to capture stochastic effects such as time-varying congestion, public transportation delays, and variable customer demand. Third, we would like to combine our model with a power-in-the-loop AMoD model [18], in order to investigate to which extent intermodality can improve the interaction with the electric power grid. Finally, we would like to explore more human-centered optimization objectives such as travel comfort and switch-over costs.

Acknowledgments

We would like to thank Dr. Daniele Vigo and Dr. Guido Gentile for the fruitful discussions, Dejan Milojevic for providing the dataset for Berlin, and Dr. Ilse New for her assistance with the proofreading and useful advice. The first author would

like to thank Dr. Chris Onder for his support. This research was supported by the National Science Foundation under CAREER Award CMMI-1454737 and the Toyota Research Institute (TRI). This article solely reflects the opinions and conclusions of its authors and not NSF, TRI, or any other entity. This paper is dedicated to Stella.

APPENDIX

A. Elements of Graph Theory

In the following discussions, we use the following definitions, adapted from [61].

Definition A.1 (Directed Graph and Directed Path)**.** *A directed graph (digraph) of order n is a pair $\mathcal{G} = (\mathcal{V}, \mathcal{A})$, with a set $\mathcal{V}$ of n elements (vertices) and a set $\mathcal{A} \subseteq \mathcal{V} \times \mathcal{V}$ of ordered pairs of vertices (arcs). A directed path is an ordered sequence of vertices such that any pair of consecutive vertices represents an arc in the digraph.*

Definition A.2 (Strongly Connected Graph)**.** *A digraph $\mathcal{G}$ is strongly connected if there exists a directed path between any two vertices.*

Definition A.3 (Incidence Matrix)**.** *Let $\mathcal{G}$ be a directed graph with n nodes and m arcs. Number the nodes of $\mathcal{G}$ with a unique $i \in \{1,\dots,n\}$ and its arcs with a unique $a \in \{1,\dots,m\}$. The incidence matrix $B\{-1,0,1\}^{n\times m}$ of $\mathcal{G}$ is defined as*

$$[B]_{ia} = \begin{cases} +1 & \text{if node } i \text{ is the source node of arc } a, \\ -1 & \text{if node } i \text{ is the sink node of arc } a, \\ 0 & \text{otherwise.} \end{cases} \tag{26}$$

B. Proof of Theorem 3.1

Proof. The proof relies on showing that the prices and tolls (22)–(24) align the incentives of self-interested agents with the social optimum. Namely, customers account for the induced congestion and the AMoD operational costs through the road prices, and for the public transit operation through its fares. Similarly, the road tolls ensure that a selfish AMoD system operator accounts for the congestion caused by its vehicles.

By Lemmas II.2, III.1, and III.2 the proof of 1) and 2) reduces to showing that if $(\{x_m\}_m^\star, x_0^\star)$ is a solution to Problem 2, then $x_m^\star$ is a solution to Problem 5 for all $m \in \mathcal{M}$ and $x_0^\star$ is a solution to Problem 6. Part 3) follows as a corollary. Informally, we proceed by showing that under the given prices if $(\{x_m\}_m^\star, x_0^\star)$ satisfies the KKT conditions for Problem 2, then $x_m^\star$ and $x_0^\star$ satisfies the KKT conditions for their respective problems. As all problems are strictly convex and satisfy the linearity constraint qualification, the KKT conditions provide necessary and sufficient conditions for optimality [62].

Formally, we first prove 1). To simplify the notation let $G_m(\lambda, x) \coloneqq \lambda^\top (Bx - b_m)$, $\mathcal{N} \coloneqq \{1,\dots,|\mathcal{A}|\}$ and $\mathcal{N}_\mathrm{R} \coloneqq \{1,\dots,|\mathcal{A}_\mathrm{R}|\}$. The KKT conditions suggest that $(\{x_m\}_m^\star, x_0^\star)$ is a solution to Problem 2 iff there exists $(\{\lambda_{\mathrm{C},m}, \mu_{\mathrm{NN},m}\}_m, \mu_{\mathrm{NN},0}, \lambda_\mathrm{R}, \mu_\mathrm{cR})$ so that

$$\begin{aligned}
0 &= V_\mathrm{Q} \cdot x_{\mathrm{R},m}^\star + c_\mathrm{R} + \frac{\partial G_m(\lambda_{\mathrm{C},m}, x_m^\star)}{\partial x_{\mathrm{R},m}} - \mu_{\mathrm{NN,R},m} \\
&\quad + B_R^\top \lambda_\mathrm{R} + \mu_\mathrm{cR} && \forall m \in \mathcal{M} \\
0 &= V_\mathrm{Q} \cdot x_{\mathrm{W},m}^\star + c_\mathrm{W} + \frac{\partial G_m(\lambda_{\mathrm{C},m}, x_m^\star)}{x_{\mathrm{W},m}} - \mu_{\mathrm{NN,W},m} && \forall m \in \mathcal{M} \\
0 &= V_\mathrm{Q} \cdot x_{\mathrm{P},m}^\star + c_\mathrm{S} + \frac{\partial G_m(\lambda_{\mathrm{C},m}, x_m^\star)}{x_{\mathrm{P},m}} - \mu_{\mathrm{NN,P},m} && \forall m \in \mathcal{M} \\
0 &= V_\mathrm{Q} \cdot x_{\mathrm{C},m}^\star + c_\mathrm{C} + \frac{\partial G_m(\lambda_{\mathrm{C},m}, x_m^\star)}{x_{\mathrm{C},m}} - \mu_{\mathrm{NN,C},m} && \forall m \in \mathcal{M} \\
0 &= V_\mathrm{Q} \cdot x_0^\star + c_0 - \mu_{\mathrm{NN},0} + B_R^\top \lambda_\mathrm{R} + \mu_\mathrm{cR} \\
0 &= [\mu_{\mathrm{NN},m}]_i [x_m^\star]_i, && \forall i \in \mathcal{N}, m \in \mathcal{M} \\
0 &= [\mu_{\mathrm{NN},0}]_i [x_0^\star]_i, && \forall i \in \mathcal{N}_\mathrm{R} \\
0 &= [\mu_\mathrm{cR}]_i \left[\sum_{m\in\mathcal{M}} x_{\mathrm{R},m}^\star + x_0^\star - h_\mathrm{R} \right]_i && \forall i \in \mathcal{N}_\mathrm{R} \\
0 &\le \mu_\mathrm{cR},\ 0 \le \mu_{\mathrm{NN},m},\ 0 \le \mu_{\mathrm{NN},0},\ 0 \le x_m^\star,\ 0 \le x_0^\star && \forall m \in \mathcal{M}.
\end{aligned}$$

Similarly, $\tilde{x}_m^\star$ is a solution to Problem 5 iff there exists $(\tilde{\lambda}_{\mathrm{C},m}, \tilde{\mu}_{\mathrm{NN},m})$ such that

$$0 = V_\mathrm{Q} \cdot \tilde{x}_{\mathrm{R},m}^\star + \tilde{c}_\mathrm{R} + \frac{\partial G_m(\tilde{\lambda}_{\mathrm{C},m}, \tilde{x}_m^\star)}{\partial \tilde{x}_{\mathrm{R},m}} - \tilde{\mu}_{\mathrm{NN,R},m} \tag{27a}$$

$$0 = V_\mathrm{Q} \cdot \tilde{x}_{\mathrm{W},m}^\star + \tilde{c}_\mathrm{W} + \frac{\partial G_m(\tilde{\lambda}_{\mathrm{C},m}, \tilde{x}_m^\star)}{\partial \tilde{x}_{\mathrm{W},m}} - \tilde{\mu}_{\mathrm{NN,W},m} \tag{27b}$$

$$0 = V_\mathrm{Q} \cdot \tilde{x}_{\mathrm{P},m}^\star + \tilde{c}_\mathrm{S} + \frac{\partial G_m(\tilde{\lambda}_{\mathrm{C},m}, \tilde{x}_m^\star)}{\partial \tilde{x}_{\mathrm{P},m}} - \tilde{\mu}_{\mathrm{NN,P},m} \tag{27c}$$

$$0 = V_\mathrm{Q} \cdot \tilde{x}_{\mathrm{C},m}^\star + \tilde{c}_\mathrm{C} + \frac{\partial G_m(\tilde{\lambda}_{\mathrm{C},m}, \tilde{x}_m^\star)}{\partial \tilde{x}_{\mathrm{C},m}} - \tilde{\mu}_{\mathrm{NN,C},m} \tag{27d}$$

$$0 = [\tilde{\mu}_\mathrm{NN}]_i [\tilde{x}_m^\star]_i \qquad \forall i \in \mathcal{N} \tag{27e}$$

$$0 \le \tilde{\mu}_{\mathrm{NN},m},\ 0 \le \tilde{x}_m^\star. \tag{27f}$$

By direct comparison we observe that if $\tilde{c}_\mathrm{R} = c_\mathrm{R} + B_R^\top \lambda_\mathrm{R} + \mu_\mathrm{cR}, \tilde{c}_\mathrm{W} = c_\mathrm{W}, \tilde{c}_\mathrm{S} = c_\mathrm{S}$, and $\tilde{c}_\mathrm{C} = c_\mathrm{C}$ then $(\tilde{x}_m^\star, \tilde{\lambda}_{\mathrm{C},m}, \tilde{\mu}_{\mathrm{NN},m}) = (x_m^\star, \lambda_{\mathrm{C},m}, \mu_{\mathrm{NN},m})$ is indeed solution to (27) and $\tilde{x}_m^\star$ is therefore the unique optimal solution to Problem 5. By definition of $\tilde{c}_\mathrm{R}$ and c_S we may equivalently write

$$\begin{aligned}
p_\mathrm{R}(i,j) &= V_{\mathrm{D,R}} \cdot s_{ij} + V_\mathrm{E} \cdot e_{ij} + [B_R^\top \lambda_\mathrm{R}]_a + [\mu_\mathrm{cR}]_a \\
&= V_{\mathrm{D,R}} \cdot s_{ij} + V_\mathrm{E} \cdot e_{ij} + \lambda_\mathrm{R}(i) - \lambda_\mathrm{R}(j) + \mu_\mathrm{cR}(i,j) \\
p_\mathrm{P}(i,j) &= V_{\mathrm{D,P}} \cdot s_{ij}.
\end{aligned}$$

for all $(i,j) \in \mathcal{A}_\mathrm{R}$ and $(i,j) \in \mathcal{A}_\mathrm{P}$, respectively.

We now focus on 2): $\bar{x}_0^\star$ is a solution to Problem 6 iff there exists $(\bar{\lambda}_\mathrm{R}, \bar{\mu}_{\mathrm{NN},0})$ such that

$$0 = V_\mathrm{Q} \cdot \bar{x}_0^\star + \bar{c}_0 + B_\mathrm{R}^\top \bar{\lambda}_\mathrm{R} - \bar{\mu}_{\mathrm{NN},0} \tag{28a}$$

$$0 = [\bar{\mu}_\mathrm{NN}]_i [\bar{x}_0^\star]_i \qquad \forall i \in \mathcal{N}_\mathrm{R} \tag{28b}$$

$$0 \le \bar{\mu}_{\mathrm{NN},0},\ 0 \le \bar{x}_0^\star. \tag{28c}$$

We observe that if $\bar{c}_0 = c_0 + \mu_\mathrm{cR}$, then $(\bar{x}_0^\star, \bar{\lambda}_\mathrm{R}, \bar{\mu}_{\mathrm{NN},0}) = (x_0, \lambda_\mathrm{R}, \mu_{\mathrm{NN},0})$ solves (28) and $\bar{x}_0^\star$ is therefore the unique optimal solution to Problem 6. Hence, the pricing scheme is

$$\tau_\mathrm{R}(i,j) = [\mu_\mathrm{cR}]_a = \mu_\mathrm{cR}(i,j) \quad \forall (i,j) \in \mathcal{A}_\mathrm{R}.$$

Finally, we focus on 3) and show that, up a regularization term, AMoD operators make no profit. Their revenue equals incomes minus costs:

$$\begin{aligned}
r &= (B_\mathrm{R}^\top \lambda_\mathrm{R})^\top \sum_{m\in\mathcal{M}} x_m - \bar{c}_0^\top x_0 = -(B_\mathrm{R}^\top \lambda_\mathrm{R} + \bar{c}_0)^\top x_0 \\
&= -(B_\mathrm{R}^\top \lambda_\mathrm{R} + \bar{c}_0 - \mu_{\mathrm{NN},0})^\top x_0 = V_\mathrm{Q} \cdot x_0^\top x_0,
\end{aligned}$$

which is indeed zero up to regularization.

This concludes the proof. □

REFERENCES

[1] J. I. Levy, J. J. Buonocore, and K. Von Stackelberg, "Evaluation of the public health impacts of traffic congestion: a health risk assessment," *Environmental Health*, vol. 9, no. 1, p. 65, 2010.

[2] B. Tuttle and T. Cowles, "Traffic jams cost americans $124 billion in 2013," *Time - Money*, 2014.

[3] EPA, "Inventory of us greenhouse gas emissions and sinks: 1990-2016," *Environmental Protection Agency*, 2018.

[4] R. Cervero, A. Golub, and B. Nee, "City carshare: longer-term travel demand and car ownership impacts," *Transportation Research Record: Journal of the Transportation Research Board*, no. 1992, pp. 70–80, 2007.

[5] R. R. Clewlow and G. S. Mishra, "Disruptive transportation: The adoption, utilization, and impacts of ride-hailing in the United States," Inst. of Transportation Studies, Univ. of California, Davis, Tech. Rep. UCD-ITS-RR-17-07, 2017.

[6] W. Hu. (2017) Your Uber car creates congestion. should you pay a fee to ride? The New York Times. Available online.

[7] R. Zhang and M. Pavone, "Control of robotic Mobility-on-Demand systems: A queueing-theoretical perspective," *Int. Journal of Robotics Research*, vol. 35, no. 1–3, pp. 186–203, 2016.

[8] S. Banerjee, R. Johari, and C. Riquelme, "Pricing in ride-sharing platforms: A queueing-theoretic approach," in *ACM Conf. on Economics and Computation*, 2015.

[9] R. Iglesias, F. Rossi, R. Zhang, and M. Pavone, "A BCMP network approach to modeling and controlling autonomous mobility-on-demand systems," *Int. Journal of Robotics Research*, vol. 38, no. 2-3, p. 357?374, 2019.

[10] 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.

[11] M. Maciejewski, J. Bischoff, S. Hörl, and K. Nagel, "Towards a testbed for dynamic vehicle routing algorithms," in *Int. Conf. on Practical Applications of Agents and Multi-Agent Systems - Workshop on the application of agents to passenger transport (PAAMS-TAAPS)*, 2017.

[12] S. Hörl, C. Ruch, F. Becker, E. Frazzoli, and K. W. Axhausen, "Fleet control algorithms for automated mobility: A simulation assessment for Zurich," in *Annual Meeting of the Transportation Research Board*, 2018.

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

[14] K. Spieser, K. Treleaven, R. Zhang, E. Frazzoli, D. Morton, and M. Pavone, "Toward a systematic approach to the design and evaluation of Autonomous Mobility-on-Demand systems: A case study in Singapore," in *Road Vehicle Automation*. Springer, 2014.

[15] R. Iglesias, F. Rossi, K. Wang, D. Hallac, J. Leskovec, and M. Pavone, "Data-driven model predictive control of autonomous mobility-on-demand systems," in *Proc. IEEE Conf. on Robotics and Automation*, 2018.

[16] M. Salazar, M. Tsao, I. Aguiar, M. Schiffer, and M. Pavone, "A congestion-aware routing scheme for autonomous mobility-on-demand systems," in *European Control Conference*, 2019.

[17] K. Solovey, M. Salazar, and M. Pavone, "Scalable and congestion-aware routing for autonomous mobility-on-demand via frank-wolfe optimization," in *Robotics: Science and Systems*, 2019, submitted.

[18] F. Rossi, R. Iglesias, M. Alizadeh, and M. Pavone, "On the interaction between Autonomous Mobility-on-Demand systems and the power network: Models and coordination algorithms," in *Robotics: Science and Systems*, 2018, Extended version available at https://arxiv.org/abs/1709.04906.

[19] M. Pavone, S. L. Smith, E. Frazzoli, and D. Rus, "Robotic load balancing for Mobility-on-Demand systems," *Int. Journal of Robotics Research*, vol. 31, no. 7, pp. 839–854, 2012.

[20] I. Mayeres and S. Proost, "Congested roads and general equilibrium Pigouvian tax solutions," in *Recent Advances in Spatial Equilibrium Modelling*. Springer, 1996, pp. 221–242.

[21] H. Yang and M. G. Bell, "Traffic restraint, road pricing and network equilibrium," *Transportation Research Part B: Methodological*, vol. 31, no. 4, pp. 303 – 314, 1997.

[22] H. Yang and H.-J. Juang, *Mathematical and Economic Theory of Road Pricing*. Elsevier, 2005.

[23] A. de Palma and R. Lindsey, "Traffic congestion pricing methodologies and technologies," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 6, pp. 1377 – 1399, 2011.

[24] D. Brownstone and K. A. Small, "Valuing time and reliability: assessing the evidence from road pricing demonstrations," *Transportation Research Part A: Policy and Practice*, vol. 39, no. 4, pp. 279–293, 2005.

[25] J. Leape, "The london congestion charge," *Journal of Economic Perspectives*, vol. 20, no. 4, pp. 157 – 176, 2006.

[26] J. Eliasson, L. Hultkrantz, L. Nerhagen, and L. S. Rosqvist, "The stockholm congestion – charging trial 2006: Overview of effects," *Transportation Research Part A: Policy and Practice*, vol. 43, no. 3, pp. 240 – 250, 2009.

[27] M. D. Simoni, K. M. Kockelman, K. M. Gurumurthy, and J. Bischoff, "Congestion pricing in a world of self-driving vehicles: an analysis of different strategies in alternative future scenarios," *arXiv preprint arXiv:1803.10872*, 2018.

[28] T. D. Chen and K. M. Kockelman, "Management of a shared autonomous electric vehicle fleet: Implications of pricing schemes," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2572, pp. 37–46, 2016.

[29] A. Vakayil, W. Gruel, and S. Samaranayake, "Integrating shared-vehicle Mobility-on-Demand systems with public transit," in *Annual Meeting of the Transportation Research Board*, 2017.

[30] G. Gentile and K. Noekel, Eds., *Modelling Public Transport Passenger Flows in the Era of Intelligent Transport Systems*. Springer New York, 2016.

[31] J. Bischoff, I. Kaddoura, M. Maciejewski, and K. Nagel, "Re-defining the role of public transport in a world of shared autonomous vehicles," in *Symposium of the European Association for Research in Transportation (hEART)*, 2017.

[32] C. Seaborn, J. Attanucci, and N. Wilson, "Analyzing multimodal public transport journeys in london with smart card fare payment data," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2121, pp. 55–62, 2009.

[33] W. Hu and V. Wang, "Congestion plan for Manhattan gets mixed reviews," *The New York Times*, 2018.

[34] M. Salazar, F. Rossi, M. Schiffer, C. H. Onder, and M. Pavone, "On the interaction between autonomous mobility-on-demand and the public transportation systems," in *Proc. IEEE Int. Conf. on Intelligent Transportation Systems*. IEEE, 2018, pp. 2262–2269, best Student Paper. Extended Version, Available at https://arxiv.org/abs/1804.11278.

[35] Bureau of Public Roads, "Traffic assignment manual," U.S. Dept. of Commerce, Urban Planning Division, Tech. Rep., 1964.

[36] W. Reilly, *Highway Capacity Manual 2000*. Transportation Research Board of the National Academies, 1997.

[37] L. Guzzella and A. Sciarretta, *Vehicle Propulsion Systems*. Springer Berlin Heidelberg, 2007.

[38] (2018) Nrel drivecat - chassis dynamometer drive cycles - new york city cycle. National Renewable Energy Laboratory. Available online. [Online]. Available: http://www.nrel.gov/transportation/drive-cycle-tool

[39] L. Guzzella and C. H. Onder, *Introduction to Modeling and Control of Internal Combustion Engine Systems*, second edition ed. Springer Berlin Heidelberg, 2004.

[40] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic Theory*. Oxford Univ. Press, 1995.

[41] H. Neuburger, "The economics of heavily congested roads," *Transportation Research*, vol. 5, no. 4, pp. 283–293, 1971.

[42] F. Rossi, "On the interaction between Autonomous Mobility-on-Demand systems and the built environment: Models and large scale coordination algorithms," Ph.D. dissertation, Stanford University, Dept. of Aeronautics and Astronautics, 2018.

[43] M. Tsao, D. Milojevic, C. Ruch, M. Salazar, E. Frazzoli, and M. Pavone, "Model predictive control of ride-sharing autonomous mobility on demand systems," in *Proc. IEEE Conf. on Robotics and Automation*, 2019, in Press.

[44] G. LLC. (2018) Google maps for iPhone. Available at https://sites.google.com/a/pressatgoogle.com/google-maps-for-iphone/google-maps-metrics.

[45] R. Sugar. (2017) Uber and Lyft cars now outnumber yellow cabs in NYC 4 to 1. Curbed. Vox Media, Inc. Available at https://ny.curbed.com/2017/10/13/16468716/uber-yellow-cab-nyc-surpass-ridership.

[46] A. Horni, K. Nagel, and K. W. Axhausen, Eds., *The Multi-Agent Transport Simulation MATSim*. Ubiquity Press, 2016.

[47] D. Ziemke and K. Nagel, "Development of a fully synthetic and open scenario for agent-based transport simulations -– the matsim open berlin scenario," Transport Systems Planning and Transport Telematics – Technische Universität Berlin, Tech. Rep., 2017.

[48] M. Haklay and P. Weber, "OpenStreetMap: User-generated street maps," *IEEE Pervasive Computing*, vol. 7, no. 4, pp. 12–18, 2008.

[49] New York City Open Data. (2018) New York City Subway Lines. Available at https://data.cityofnewyork.us/Transportation/Subway-Lines/3qz8-muuul.
[50] MTA. (2018) MTA Subway Schedules. Available at http://www.mta.info/schedules.
[51] B. O. Data. Vbb-fahrplandaten via gtfs. Berlin Partner and BerlinOnline Stadtportal GmbH & Co. Available online. [Online]. Available: https://daten.berlin.de/datensaetze/vbb-fahrplandaten-gtfs
[52] P. Mosendz and H. Sender, "Exclusive: Here's how long it takes to get an UBER in U.S. cities," *Newsweek*, 2014.
[53] W. Time. (2018, Mar.) Carboon footprint data. Wired. Available at https://api.watttime.org. Retrieved on March 23, 2018.
[54] Statista. Entwicklung des co2-emissionsfaktors für den strommix in deutschland in den jahren 1990 bis 2017 (in gramm pro kilowattstunde). Statista GmbH. Available online. [Online]. Available: https://de.statista.com/statistik/daten/studie/38897/umfrage/co2-emissionsfaktor-fuer-den-strommix-in-deutschland-seit-1990/
[55] U.S. Dept. of Transportation, "Revised departmental guidance on valuation of travel time in economic analysis," Tech. Rep., 2015.
[56] Bureau of Transportation Statistics, "National transportation statistics." U.S. Dept. of Transportation, Tech. Rep., 2016.
[57] J. Neff and M. Dickens, "2016 public transportation fact book," American Public Transportation Association, Tech. Rep., 2017.
[58] A. Levy, "Why are the nyc subway's operating costs so high?" *Curbed*, 2017, available online.
[59] E. S. Rueb, "How New York City gets its electricity," *The New York Times*, 2017.
[60] B. Wehrmann, "Germany's power price reaches new all-time high in march 2019, gas prices spike too," *Journalism for the Energy Transition*, 2019, available online.
[61] F. Bullo, *Lectures on Network Systems*. CreateSpace, 2018.
[62] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge Univ. Press, 2004.