

Analysis and Control of Autonomous Mobility-on-Demand Systems

Gioele Zardini,^{1,*} Nicolas Lanzetti,^{2,*} Marco Pavone,³
and Emilio Frazzoli¹

¹Institute for Dynamic Systems and Control, Department of Mechanical and Process Engineering, ETH Zürich, Zürich, Switzerland; email: gzardini@ethz.ch, efrazzoli@ethz.ch

²Automatic Control Laboratory, Department of Information Technology and Electrical Engineering, ETH Zürich, Zürich, Switzerland; email: lnicholas@ethz.ch

³Autonomous Systems Lab, Department of Aeronautics and Astronautics, Stanford University, Stanford, California, USA; email: pavone@stanford.edu

ANNUAL
REVIEWS **CONNECT**

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Annu. Rev. Control Robot. Auton. Syst. 2022.
5:633–58

First published as a Review in Advance on
November 23, 2021

The *Annual Review of Control, Robotics, and
Autonomous Systems* is online at
control.annualreviews.org

<https://doi.org/10.1146/annurev-control-042920-012811>

Copyright © 2022 by Annual Reviews.
All rights reserved

*These authors contributed equally to this article

Keywords

autonomous vehicles, autonomous mobility-on-demand, fleet management, ride-sharing, future mobility systems

Abstract

Challenged by urbanization and increasing travel needs, existing transportation systems need new mobility paradigms. In this article, we present the emerging concept of autonomous mobility-on-demand, whereby centrally orchestrated fleets of autonomous vehicles provide mobility service to customers. We provide a comprehensive review of methods and tools to model and solve problems related to autonomous mobility-on-demand systems. Specifically, we first identify problem settings for their analysis and control, from both operational and planning perspectives. We then review modeling aspects, including transportation networks, transportation demand, congestion, operational constraints, and interactions with existing infrastructure. Thereafter, we provide a systematic analysis of existing solution methods and performance metrics, highlighting trends and trade-offs. Finally, we present various directions for further research.

1. INTRODUCTION

Countries worldwide are undergoing dramatic urbanization: Today, 55% of the world's population lives in urban areas, and the proportion is expected to reach 68% by 2050 (1). Urbanization leads to an increase in urban travel, and hence also in the externalities that this travel produces. Indeed, cities are considered responsible for 78% of the world's energy consumption and more than 60% of global emissions (30% of which is produced by transportation, in the United States) (2). Furthermore, congestion levels are increasing worldwide: In the United States, congestion causes 8.8 billion hours of extra travel time, equivalent to more than a week of vacation for the average US commuter, causing \$180 billion of delay and fuel costs, which is close to 1% of the US gross domestic product (3).

In this context, private cars are considered an unsustainable solution for future personal mobility (4). Indeed, vehicles are typically overengineered (designed for long-distance highway travel but used mostly for low-speed urban trips) and underutilized (private cars are usually parked more than 90% of the time). The mobility-on-demand (MoD) concept has potential to alleviate these issues. In a MoD system, (light, electric) vehicles are placed throughout a city, and when customers want to reach a certain destination, they can walk to the nearest vehicle, pick it up, drive to the destination, and drop it off. MoD services directly target issues such as pollution and parking spaces (due to high utilization rates) and satisfy needs of personal mobility (4, 5).

However, MoD systems suffer from limitations, such as their tendency to become unbalanced due to unevenly distributed transportation demand. To overcome such issues, ride-hailing companies started offering a variation of MoD services, whereby passengers are assigned by a central manager to an on-demand driver, who then drives them to their destination. Although such services have become very popular [in New York City, ride-hailing companies increased their number of rides by 1,000% between 2012 and 2019 (6)], they exploit public resources (e.g., roads), require a large number of drivers to scale, and often lead to potentially disruptive consequences for the transportation system and for society at large (7, 8). Furthermore, their profitability is questionable.

Technological advances in the field of autonomous driving over the past few decades offer a new mobility paradigm: autonomous mobility-on-demand (AMoD). An AMoD system consists of a fleet of autonomous vehicles (AVs) that pick up passengers and transport them to their destination. A manager controls the fleet by simultaneously assigning passengers to AVs, routing them, and rebalancing the fleet by relocating customer-free AVs to realign their geographical distribution with transportation demand. One could argue that MoD systems in which drivers precisely follow instructions from a fleet manager could provide the same service as AMoD systems. So what justifies AMoD over MoD?

AMoD promises two main benefits. First, it increases the supply of drivers to match increasing demand with good quality of service. Indeed, American drivers spend approximately 1 h per day driving (9), equivalent to one-ninth of a work day, considering commute time to be productive time. Assuming driving is allocated to the same group of people (i.e., ignoring demand peaks), one out of nine workers should be a MoD driver. Additionally, customer-free trips, for which a common lower bound is 30% (e.g., based on Singapore data), increase the proportion to one out of seven, hindering the scalability of MoD systems. Second, AMoD drastically reduces transportation costs (10, 11). For instance, Becker et al. (11) estimated that electrification and automation would reduce the costs of taxis by more than 70% in Austin and 85% in Zürich.

Analytically and computationally, the operation of AMoD systems is similar to that of taxi systems, with one key difference: A taxi system operator cannot enforce assignments of passengers to drivers and balanced vehicle distributions because drivers have the final say. Typically, taxi drivers

act selfishly to maximize their own profit, while AVs can act cooperatively. In AMoD systems, central coordination schemes can minimize externalities by optimizing routing, rebalancing, charging schedules, and so on. Furthermore, customers do not need to drive and can thus better use commute time. Driven by such promising benefits, various companies started to develop the technology (approximately 30 AV companies just in California), and some are already providing mobility services. Examples include Motional, whose AVs have provided more than 100,000 rides in Las Vegas (12), and Waymo, whose AVs have driven more than 20 million miles (13).

Overall, the novelty of AMoD systems calls for methods for analysis and control. The challenges are sociotechnical, ranging from technical ones (e.g., reaching sufficient levels of autonomy and defining operational regulations and liabilities) to political and economic ones (e.g., taxation policies, privacy, and fairness) (14). Questions of interest include which kind of autonomy and what levels of safety and reliability are needed for AVs to operate in cities, how many AVs are needed to guarantee a desired quality of service, how AMoD systems will interact with existing mobility systems, whether AMoD systems will reduce congestion, and whether we can use AMoD systems to steer urban mobility toward a sustainable future.

Numerous reviews have been published on aspects relevant to the analysis and control of AMoD systems, ranging from the vehicle routing problem (15–20) to shared (autonomous) mobility (21–27). However, to the best of our knowledge, no review has focused specifically on the analysis and control of AMoD systems, identifying problem settings, models, solution approaches, and future challenges.

In this article, we provide a synoptic review of methods for the analysis and control of AMoD systems. Specifically, our contribution comprises four parts. First, we identify problem settings for analysis and control of AMoD systems. Second, we present mathematical models, including models for the transportation network, transportation demand, congestion, operational constraints, and interactions with existing infrastructure. Third, we review solution methods, together with performance metrics and case studies. Our analysis highlights patterns and instances yet to be solved. Finally, we provide an outlook on future research.

2. ANALYSIS AND CONTROL OF AUTONOMOUS MOBILITY-ON-DEMAND SYSTEMS

Before introducing mathematical tools to reason about AMoD systems, we clarify why we need to analyze and control them in the first place. Essentially, there are two aspects:

- **Operational policies:** The deployment of AMoD systems calls for real-time operational policies. Questions of interest include which AV should pick up a customer requesting a ride, how AVs should be routed throughout the city, what customer-free AVs should do, and whether multiple requests can be simultaneously accommodated by sequentially picking up customers and driving them to their destinations.
- **Planning studies:** Operations of AMoD systems rely on planning. Questions to answer include how many AVs are needed to provide a given quality of service, where charging stations should be located, and how AVs will affect existing mobility systems.

Addressing these questions requires two ingredients: a mathematical model of the AMoD system and a solution approach (**Figure 1**). We first review problems related to the analysis and control of AMoD systems, distinguishing operational policies (Section 2.1) and planning (Section 2.2). In Section 3, we investigate models for AMoD systems, describing models for transportation networks, transportation demand, congestion, operational constraints, and interactions. Finally, in

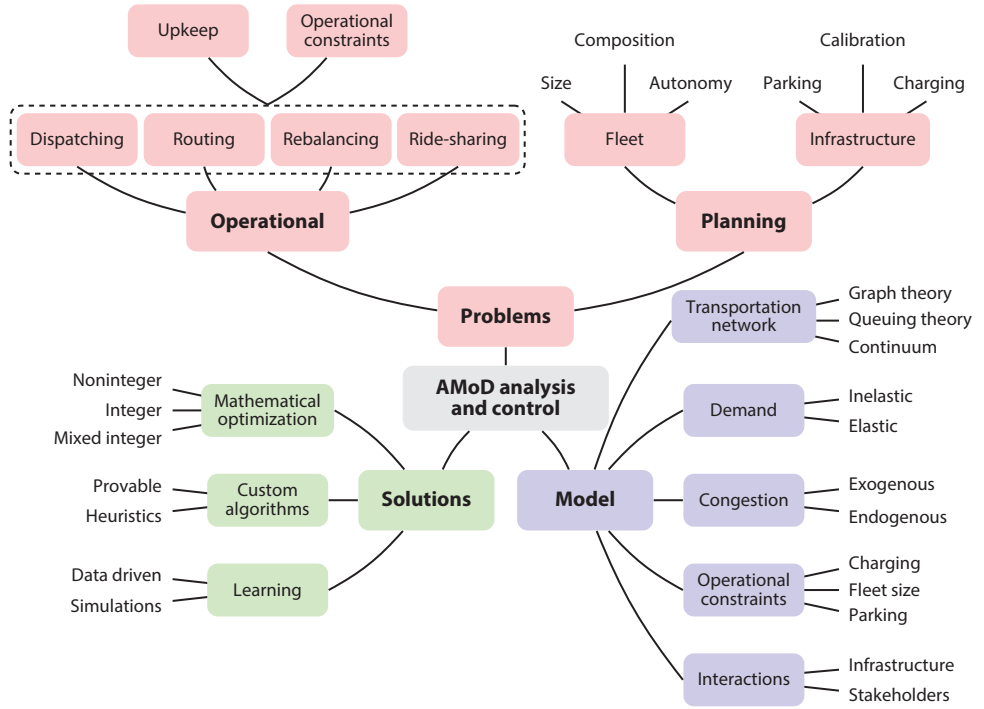


Figure 1

Dependency diagram for the analysis and control of autonomous mobility-on-demand (AMoD) systems, featuring mathematical models, problems, and solution approaches.

Section 4, we present solution approaches for operational matters and planning problems as well as related performance metrics.

2.1. Operational Policies

We distinguish four operational tasks: dispatching, routing, rebalancing, and ride-sharing (**Figure 2**). Additionally, each task could account for the upkeep of AVs, which includes scheduling of charging, calibration, parking, and general maintenance routines.

2.1.1. Dispatching. Given a set of customer requests, the fleet manager needs to assign them to available AVs (**Figure 2a**). At their core, dispatching problems are assignment problems, which have been well studied in operations research (28–31). Typically, when solving such problems, one seeks a system optimum that maximizes fleet performance.

2.1.2. Routing. After being dispatched, AVs need to be (dynamically) routed to the assigned customers and then to the destinations of their trips (**Figure 2b**). Routes are typically chosen to optimize specific metrics (e.g., travel time and travel distance), possibly considering road congestion. Such problems belong to the family of vehicle routing problems, in particular dynamic pickup and delivery problems (reviewed in 18, 32–34).

2.1.3. Rebalancing. Rebalancing problems deal with repositioning of customer-free AVs (35) (**Figure 2c**). Policies aim to maximize the responsiveness of AMoD systems to new requests while minimizing the imbalance of the AV fleet caused by asymmetries in transportation demand.

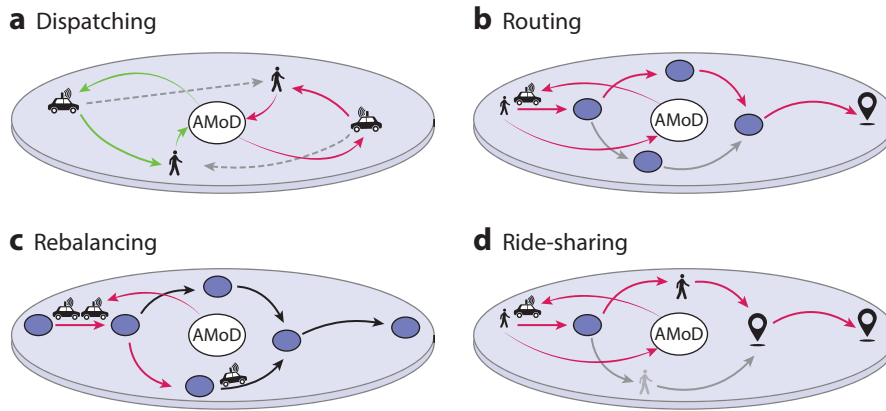


Figure 2

Operational policies for autonomous mobility-on-demand (AMoD) systems. An AMoD fleet manager gathers information from customers and autonomous vehicles, processes it, and produces actionable policies for the vehicles.

Rebalancing techniques usually consist of two steps: an estimate of future transportation demand and an efficient redistribution of customer-free AVs.

2.1.4. Ride-sharing. Ride-sharing refers to the process of simultaneously serving multiple, unrelated customer requests with a single AV (**Figure 2d**). This additional feature complicates the solution of the aforementioned problems: Ride-sharing policies need to provide AV assignments and routes dealing with multiple pickup and drop-off locations while maximizing fleet performance and not causing a deterioration in quality of service (36–38).

2.2. Planning Autonomous Mobility-on-Demand Operations

Numerous planning problems have been investigated, which we cluster in three categories.

2.2.1. Design of autonomous vehicle fleets. The design of AV fleets impacts the performance of AMoD systems. Design problems include fleet sizing (39–43), fleet composition (e.g., vehicle capacity and propulsion) (44), and autonomy (45, 46). For instance, Spieser et al. (39) solved two instances of fleet sizing: minimum fleet sizing and performance-driven fleet sizing. The former seeks the minimum number of AVs in a fleet to keep the number of unmet demands bounded [92,693 AVs in Singapore (39)], while the latter quantifies the number of AVs for a certain quality of service. Fleet sizing also informs auxiliary operations related to AMoD systems. For instance, in low-demand regimes, unnecessary AVs could be sent to charging and calibration facilities or provide auxiliary services, such as package delivery (47).

2.2.2. Interactions with existing infrastructure. AMoD systems operate while using public and private resources, such as space for parking and idling, roads, and electricity. Therefore, planning studies also focus on interactions between AMoD systems and existing infrastructure. Aspects of interest include the allocation of parking space (48, 49), the optimal placement of charging stations, charge scheduling for electric AVs (50–54), contributions to congestion (55), joint operation with other modes of transportation (56, 57), autonomous calibration (58), smart infrastructure (59), and the achievable performance of the overall mobility system (45).

2.2.3. Interactions among stakeholders. AMoD systems are part of a large ecosystem of mobility providers that includes public transportation systems and private companies. Importantly, studies investigate the interplay among AMoD systems, other mobility solutions, and regulatory agencies (e.g., municipalities) (60, 61), focusing on pricing strategies, incentives, and tolls (51, 62–65). While constituting an exciting research direction, the study of interactions among stakeholders is not the focus of this article. Nevertheless, we discuss the topic when posing future challenges in Section 5.

3. MATHEMATICAL MODELING OF AUTONOMOUS MOBILITY-ON-DEMAND SYSTEMS

In this section, we review models for AMoD systems, including modeling approaches for the transportation network, AVs, transportation demand, congestion models, operational constraints, and interactions with existing infrastructure. We focus the exposition on mathematical models. We discuss simulation-based and empirical approaches, often combined with machine learning, in Section 4.

3.1. Transportation Network and Autonomous Vehicles

Models of transportation networks and AVs can be clustered in two categories:

- Time-varying models consider both the spatial and the temporal dimensions of the transportation system. Typically, authors focus on a specific geographical area (e.g., Manhattan) and a specific time window (e.g., one day). This approach is common when modeling microscopic phenomena or making operational decisions but usually leads to more complicated models (e.g., larger graphs).
- Time-invariant models assume a transportation system in steady state, where one neglects the temporal dimension, by assuming a specific time window and treating it as a single time unit (e.g., considering the number of AVs per unit time). This assumption is reasonable if requests change slowly compared with the average travel time of individual trips, as is often observed in densely populated urban environments (56, 66). This approach is popular in mesoscopic and macroscopic studies, as it simplifies the setting while capturing the main aspects of the problem.

In the following, we present mathematical models and relate them to the time-varying and time-invariant approaches (**Figure 3**).

3.1.1. Graph-theoretic models. The transportation network is modeled as a directed graph $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$, where \mathcal{V} is the set of vertices and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges (**Figure 3a**). A node $v \in \mathcal{V}$ represents a location, such as a station or an urban area, and an edge $\langle v_1, v_2 \rangle \in \mathcal{E}$ represents a road (or a combination of roads) from location v_1 to location v_2 . Each edge is usually associated with metrics such as travel time $t : \mathcal{E} \rightarrow \mathbb{R}_{\geq 0}$ and distance $d : \mathcal{E} \rightarrow \mathbb{R}_{\geq 0}$ [e.g., $t_{ij} := t(i, j)$ is the travel time of edge $\langle i, j \rangle$].

When using graph-theoretic models, one often adopts an (Eulerian) fluid-dynamic approach: AVs and customers are represented not individually but as (possibly noninteger) flows between nodes. In this setting, time-invariant models usually reduce to static network flow problems (for details, see 67). Dynamic models, by contrast, are studied via differential equations (35, 68–71) or via time-expanded graphs (72). Differential equations capture the dynamic behavior of AVs and customers on the graph. For instance, Pavone et al. (35) showed that, by flow conservation, the number of AVs $v_i(t)$ and the number of customers $c_i(t)$ at node i obey the nonlinear time-delayed

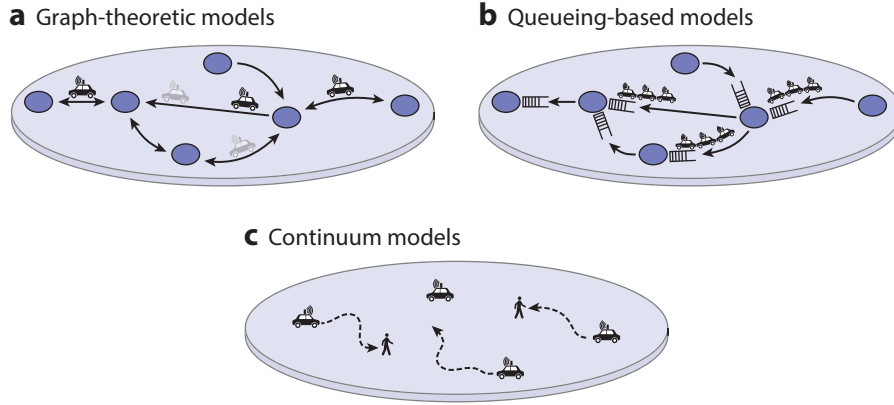


Figure 3

Models for a transportation network and autonomous vehicles.

differential equations

$$\begin{aligned}
 \dot{v}_i(t) = & -\lambda_i H(v_i(t)) + (\lambda_i - v_i) H(v_i(t)) H(c_i(t)) \\
 & + \sum_{j \neq i} p_{ij} (\lambda_j H(v_j(t - t_{ji})) - (\lambda_j - \mu_j) H(c_j(t - t_{ji})) H(v_j(t - t_{ji}))) \\
 & - \sum_{j \neq i} \alpha_{ij} H(v_i(t)) + \sum_{j \neq i} \alpha_{ji} H(v(t - t_{ji})), \\
 \dot{c}_i(t) = & \lambda_i (1 - H(v_i(t))) + (\lambda_i - v_i) H(v_i(t)) H(c_i(t)),
 \end{aligned}
 \tag{1}$$

where λ_i and μ_i are the rates of arrival and departure of customers at node i , α_{ij} is the rate of rebalancing AVs from node i to node j , p_{ij} is the probability that a customer at node i wants to travel to node j , and $H(x)$ is the Heaviside function [$H(x) = 1$ if $x > 0$, $H(x) = 0$ otherwise]. Time-expanded graphs, by contrast, are a natural generalization of directed graphs for time-varying transportation networks. Herein, nodes are spatiotemporal and consist of a geographical location and a time instant: An AV is at $n = \langle v, t \rangle$ if it is at the physical node v at time t (73). Accordingly, an edge between $n_1 = \langle v_1, t_1 \rangle$ and $n_2 = \langle v_2, t_2 \rangle$ exists if and only if v_2 can be reached from v_1 in the time interval $t_2 - t_1$.

While most works adopt a fluidic approach, graph-theoretic models can be combined with vehicle-centric models, whereby AVs and customers are modeled individually (74–76). For instance, James & Lam (75) introduced a binary decision variable $x_{ij}^k \in \{0, 1\}$ to indicate whether AV k is traversing edge $\langle i, j \rangle$ in order to optimize AV routes. Nonetheless, vehicle-centric models sometimes ignore the transportation network and model only AVs and customers (77–79). In the simplest setting, dispatching of AVs is optimized via binary decision variables $y_{ij} \in \{0, 1\}$ that take value 1 if and only if AV i is assigned to customer j .

3.1.2. Queueing-theoretic models. Queueing theory deals with studies of waiting lines. In AMoD systems, trips are modeled as queues between locations (**Figure 3b**), allowing one to analyze characteristics such as the availability of AVs (i.e., the probability that at least one AV is available), which is closely related to waiting time. Formally, consider N stations placed at specific geographical locations and m AVs providing mobility service (80–82). Customers arrive at each station according to an exogenous stochastic process and select a destination with some

probability (for more details on demand models, see Section 3.2). If AVs are available at a station, customers travel to their respective destinations; otherwise, they leave the system (so-called passenger loss). Travel times between stations are also stochastic and are usually modeled as exponentially distributed random variables. When designing policies for AMoD systems, one analyzes and regulates how AVs move from one queue to another.

These models are often deployed in time-invariant settings. The case of infinitely large fleets ($m \rightarrow \infty$) has been of particular interest to derive theoretic insights. For instance, George (80) showed that, for a fleet of m AVs, the availability $A_i(m)$ of AVs at node i satisfies

$$\lim_{m \rightarrow \infty} A_i(m) = \frac{\lambda_i / \pi_i}{\max_{i \in \{1, \dots, N\}} \lambda_i / \pi_i},$$

where λ_i is the average arrival rate at station i and π_i is the relative throughput at station i . These insights formally quantify the imbalance of the AMoD system and allow one to rigorously reason about rebalancing policies (82).

3.1.3. Continuum models. The transportation network is modeled as a compact domain $\Omega \subset \mathbb{R}^2$ of the Euclidean plane (**Figure 3c**). AVs are free to move on Ω and are typically assumed to have a constant speed. Abstracting away the topology of the road network substantially simplifies the analysis and synthesis of control strategies for AMoD systems (83), allowing one to derive analytical expressions for relevant operational parameters (18). For instance, Treleven et al. (84) proved that for particular origin–destination distributions, the distance driven by customer-free AVs is lower bounded by the earth mover’s distance (85) and provides sufficient and necessary conditions for load balancing of transportation systems. Recently, researchers also modeled AMoD systems using stochastic differential equations (86). Here, the stochasticity of the system dynamics is driven by nonstationary demand and vehicle arrival processes, allowing one to split operational problems into an average process (which can be estimated from historical data) and an unknown disturbance.

3.2. Transportation Demand

Transportation demand (henceforth referred to as demand) models the willingness of customers to travel (e.g., for business or leisure). We distinguish between inelastic and elastic demand. A demand model is inelastic if demand is independent of the operation of the AMoD system. It is elastic if demand is affected by the operation of the AMoD system.

3.2.1. Inelastic demand. Inelastic demand models can be classified as follows:

- Deterministic demand models assume a fixed (possibly time-varying) demand distribution, often derived from historical data. Such models are typically used along with graph-theoretic models, whereby demand is modeled via origin–destination pairs or via arrival and departure rates at each node. For instance, origin–destination pairs are specified by a pair of nodes $\langle o, d \rangle \in \mathcal{V} \times \mathcal{V}$ (for time-invariant settings) or by a pair of nodes and a departure time $\langle o, d, \tau \rangle \in \mathcal{V} \times \mathcal{V} \times \mathbb{R}_{\geq 0}$ (for time-varying settings) (45, 73).
- Stochastic demand models are ubiquitous in queuing-theoretic and continuum models. Often, they result in spatial (or spatiotemporal, in time-varying settings) stochastic processes. For instance, in queuing-theoretic models, customers are assumed to arrive at each station i according to a time-invariant Poisson process and select a destination j with some probability $p_{ij} \in [0, 1]$ and $\sum_{j=1}^N p_{ij} = 1$ for all i (80–82).

3.2.2. Elastic demand. In real transportation systems, demand is elastic: A more efficient mobility service could induce demand, and conversely, a less efficient mobility service could decrease it. Herein, demand is assumed to be adversely affected by travel time. For dispatching problems, assignments under elastic demand can be computed via mathematical optimization, whereby the cost function combines congestion and disutility demand (see 87–90 and references therein). As shown by this review, researchers have focused on inelastic demand models; elastic demand, thus, represents an area for future research.

3.3. Congestion

Congestion accounts for delays in transportation systems. We distinguish between two types of congestion models: exogenous and endogenous. A congestion model is exogenous if congestion is a fixed, possibly time-varying, known parameter (affecting travel time), independent of the operation of the AMoD system. Conversely, it is endogenous if it is affected by the operation of the AMoD system.

3.3.1. Exogenous congestion. Exogenous congestion models correct (free-flow) travel time to account for congestion caused by all other users on the road. Estimates often stem from historical traffic information. These models are justified for relatively small fleet sizes: A small AMoD system will not significantly increase congestion on the roads (5). Exogenous congestion leads to more tractable models of transportation systems. For instance, when making operational decisions, one does not need to consider how policies affect travel time. Yet congestion-aware operations of AMoD systems promise to mitigate congestion (5).

3.3.2. Endogenous congestion. Endogenous congestion models are found in queuing-theoretic and graph-theoretic models. In queuing-theoretic models, congestion is typically considered through capacity constraints on the queues. Graph-theoretic models offer more flexibility: Congestion can be modeled via capacity constraints on the edges (so-called threshold models) or via increased travel times. In time-invariant network flow models, capacity constraints upper bound the total flow of vehicles, including exogenous traffic u_{ij} (i.e., vehicles already on the road) and AVs f_{ij} , on edge $\langle i, j \rangle$ (e.g., $u_{ij} + f_{ij} \leq 10$ AVs per second). For increased travel times, instead, one assumes that travel time t_{ij} on edge $\langle i, j \rangle$ depends on the road usage of all users of the transportation system, including the AMoD system itself. Formally, travel delays are expressed by a nonnegative strictly increasing function $\Phi : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ mapping the total number of vehicles on an edge to a delay: Given a flow of AVs f_{ij} on edge $\langle i, j \rangle$, travel time results from

$$t_{ij} = t_{ij}^F \cdot \Phi \left(\frac{f_{ij} + u_{ij}}{c_{ij}} \right),$$

where c_{ij} is the nominal road capacity and t_{ij}^F is the free-flow travel time.¹

3.4. Operational Constraints

To analyze and control AMoD systems, one should consider the operational limits of AVs. Here, we present the constraints studied in the literature.

¹For instance, the Bureau of Public Roads function reads $\Phi(x) = 1 + 0.15x^4$ (91, 92).

3.4.1. Fleet size. Fleet size constraints account for the limited number of AVs composing a fleet. We identify three instances of fleet size constraints. First, fleet size limitations are sometimes naturally embedded in the model of the transportation system, either because it accounts for every single AV (74, 75, 79) or because it is based on queuing theory (93). For instance, a queuing-theoretic model consists of N stations and m vehicles; thus, fleet size is implicitly specified by the model of the transportation network and AVs. Second, fleet size sometimes results from the initial condition of a dynamic model of the transportation system (35, 94). For instance, Pavone et al. (35) showed that, in the setting of Equation 1, the number of AVs is constant [i.e., $\frac{d}{dt} \sum_i v_i(t) = 0$]. Hence, fleet size results from the initial condition $\sum_i v_i(0)$. Third, fleet size can be imposed via hard constraints, especially in time-invariant network flow models, where the number of AVs can be estimated by multiplying the sum of the customer and rebalancing flows with travel time (45). Finally, we remark on two aspects of these constraints: First, each type of constraint can be readily generalized to mixed fleets by imposing limitations on each vehicle class (95), and second, fleet sizing represents not only an operational constraint but also a design parameter for planning studies (43, 45).

3.4.2. Charging. Another frequent constraint relates to operational limits of AVs. Considering ongoing electrification efforts, research has focused mainly on electric AVs and, thus, on charging constraints (52). Specifically, we distinguish two modeling approaches. The first one constrains the operation of AVs, either through charging- or refueling-related delays in travel time or through upper bounds on trip duration (74). In the second one, the battery of electric AVs has a dynamic state of charge, and operational policies are designed to optimally allocate charging schedules (53, 54, 68, 76, 96–98). Finally, researchers also consider the bidirectional interplay between transportation systems and power grids (54, 62).

3.4.3. Parking. The impact of AVs on parking spaces (and vice versa) is a relevant aspect for the management of transportation systems (49, 99). Parking policies typically relocate idling AVs while not causing deterioration in metrics such as waiting times, distance traveled, and congestion. In graph-theoretic models, effects of parking are captured by additional costs when idling at a node (price of parking) (11, 100, 101) or via upper bounds on the number of AVs at a node (capacity of parking lots) (48, 69, 70, 74). In queuing-theoretic models, stations are usually assumed to have an infinite number of parking spaces (82). Extensions to capacitated queues were discussed by Zhang (81).

3.5. Interactions

Interactions with existing infrastructure, such as public transportation systems (56) and power grids (54), as well as interactions among stakeholders in mobility ecosystems (64, 65), such as municipalities and other mobility providers, are relevant modeling aspects. For instance, the analysis of intermodal mobility systems (whereby an AMoD system provides service jointly with public transport) requires modeling both the AMoD system and the public transportation network (56). Formally, the network of each mode of transportation can be modeled as in Section 3.1.1, and networks are connected via so-called switching arcs. For simplicity, consider a road network $\mathcal{G}_R = \langle \mathcal{V}_R, \mathcal{E}_R \rangle$ and a subway network $\mathcal{G}_S = \langle \mathcal{V}_S, \mathcal{E}_S \rangle$. Switching edges (representing connections between roads and subway stations) are described by a set of edges $\mathcal{E}_C \subseteq \mathcal{V}_R \times \mathcal{V}_S \cup \mathcal{V}_S \times \mathcal{V}_R$, characterized by additional properties (e.g., walking time from a road to a subway station). The resulting network is $\mathcal{G} = \langle \mathcal{V}_R \cup \mathcal{V}_S, \mathcal{E}_R \cup \mathcal{E}_S \cup \mathcal{E}_C \rangle$ (for an extensive treatment,

see 45). This model can be used to assess benefits of intermodality (56) and to devise operational policies (72).

4. SOLUTION APPROACHES

Recall the two classes of problems introduced in Section 2:

- Operational policies are designed to control AMoD systems in real time. Given a stream of requests, the system must be able to assign them to available AVs, fulfill them, and rebalance the system, while guaranteeing fleet performance.
- Planning studies address operational problems to investigate the design of AMoD systems, interactions with existing infrastructure, and interplay among stakeholders in the transportation system.

Performance metrics and optimization objectives include revenue and costs, waiting and travel time, travel distance, demand satisfaction, fleet size, and social welfare. We review solution methods (Section 4.1) and tools for simulations (Section 4.2) from both operational and planning perspectives. We classify solution approaches in **Table 1**. We consider modeling aspects of the transportation network (continuum, fluidic, queuing-theoretic, or vehicle-centric), operational constraints (fleet size, limited AV autonomy, and parking), upkeep (solutions for scheduling problems related to charging of AVs and dynamic parking of AVs), performance metrics, solution methods (Section 4.1), case studies (Section 4.2), and references.

4.1. Review of Methods

In this section, we review methods to solve problems introduced in Section 2.

4.1.1. Dispatching. Dispatching is the first problem: Given a set of customer requests, the fleet manager needs to match them to AVs. For instance, one can rely on heuristics (38, 102, 103), such as nearest neighbors: From a queue of customer requests, one assigns the nearest AV to each request in a first-come, first-served fashion. Such methods, however, generally lead to suboptimal solutions (81). More structured approaches are based on mathematical programming (see, e.g., 31 and references therein). For instance, in vehicle-centric models, whereby one dispatches m AVs to serve n customer requests, one can introduce binary decision variables $x_{ij} \in \{0, 1\}$ that indicate whether AV i is assigned to customer request j . With r_{ij} being the profit of serving customer request j with AV i , we get the integer linear program

$$\begin{aligned} & \max_{x_{ij} \in \{0,1\}} \sum_{i=1}^m \sum_{j=1}^n r_{ij} x_{ij} \\ & \text{subject to } \sum_{i=1}^m x_{ij} \leq 1 \quad \forall j \quad (\text{customers served at most once}). \end{aligned}$$

Despite being an integer linear program, the optimization problem can be solved without integrality constraints, thanks to total unimodularity. Indeed, if an integer linear program is totally unimodular,² optimal solutions are integers, even if the integrality constraint is dropped; thus, they

²A matrix is totally unimodular if the determinant of every square submatrix is 0, -1 , or $+1$. Thus, $\min_x c^\top x$ subject to $Ax = b$ is totally modular if A is totally unimodular and b is an integer vector.

Table 1 Methods for analysis and control of autonomous mobility-on-demand systems

Problems				Model ^b							Solution									
Dispatching	Routing	Rebalancing	Ride-sharing	Upkeep ^a	Operational	Network ^c	T Invariant	D Stochastic	C Deterministic	Exogenous	Endogenous	Constraints ^d	Dispatching	Routing	Rebalancing	Ride-sharing	Upkeep	Performance metrics ^f	Case studies ^g	References
✓	✓	✓	✓	—	✓	V	✓	✓	✓	✓	✓	🚗🚗🚗	ILP	ILP	LP	ILP	—	🕒, \$	NYC	38, 132
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗🚗	← ILP →	—	—	—	—	🕒, \$	SFO	73
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗🚗	OA	OA	MILP	OA	—	🕒	NYC	110, 131
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗🚗	← OA →	—	—	—	—	🕒, km	AU	102
✓	✓	✓	✓	—	✓	V	✓	✓	✓	✓	✓	🚗🚗🚗	← ILP →	—	—	—	—	🕒, km, ☆	MIT	130
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗🚗	← ILP →	—	—	—	ILP	🕒	NYC	74
✓	✓	✓	✓	—	✓	C	✓	✓	✓	✓	✓	🚗🚗🚗	← L →	—	—	—	—	\$	CDU, W	114
✓	✓	✓	✓	—	✓	V	✓	✓	✓	✓	✓	🚗🚗🚗	← OA →	—	—	—	—	🕒	SF	103
✓	✓	✓	✓	—	✓	V	✓	✓	✓	✓	✓	—	← MILP →	—	—	MILP	—	🕒, \$	NYC, TO	123, 124
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	—	← LP →	—	—	—	—	🕒, 🚗	NYC	55
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	—	← QP →	—	—	—	—	SW	NYC, BE	56
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	—	← QP →	—	—	—	—	🕒, \$	SF	94
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	← LP →	—	—	—	—	🕒, \$, ☆	NYC, HZ	122
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	← LP →	—	—	—	—	🕒	NYC	72
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	← LP →	—	—	—	LP	🕒	AR	128
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	—	← LP →	—	—	—	—	SW, \$	DFW, OC	53, 54
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	← LP →	—	—	—	—	🕒, \$	NYC	101
✓	✓	✓	✓	—	✓	V	✓	✓	✓	✓	✓	🚗🚗	← OA →	—	—	—	—	🕒, km	CHI	79
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	—	← MILP →	—	—	—	—	🚗, 🚗, \$	HZ	120
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	← LP →	—	—	—	—	☆	SF	69
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	← CO →	—	—	—	—	\$	NYC	126
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	← LP →	—	—	—	—	\$, 🚗	SG	129
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	—	← LP →	—	—	—	—	SW	NYC	92
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	← CO →	—	—	—	—	🕒	NYC	125
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	← QP →	—	—	—	—	🕒, \$	NYC	57
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	← ILP →	—	—	—	—	🕒, \$, km	SG	95
✓	✓	✓	✓	—	✓	Q	✓	✓	✓	✓	✓	🚗🚗	← MILP →	—	—	MILP	—	☆	NB, CA	76
✓	✓	✓	✓	—	✓	—	✓	✓	✓	✓	✓	🚗🚗	← LP →	—	—	—	—	🕒	NYC	170
✓	✓	✓	✓	—	✓	—	✓	✓	✓	✓	✓	🚗🚗	← L →	—	—	—	—	🕒, ☆	NYC	115
✓	✓	✓	✓	—	✓	V	✓	✓	✓	✓	✓	🚗🚗	← LP →	—	—	—	—	🕒, \$, ☆	NYC	121
✓	✓	✓	✓	—	✓	C	✓	✓	✓	✓	✓	🚗🚗	OA	OA	—	—	—	🕒	AR	84
✓	✓	✓	✓	—	✓	V	✓	✓	✓	✓	✓	—	QCILP	QCILP	—	—	QCILP	🕒, km	CO	75
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	—	L	L	—	—	—	🕒	LA	171
✓	✓	✓	✓	—	✓	Q	✓	✓	✓	✓	✓	🚗🚗	MILP	—	NLP	—	—	☆	NY, HZ	93
✓	✓	✓	✓	—	✓	C	✓	✓	✓	✓	✓	🚗🚗	OA	—	OA	—	—	🕒	SFO	83
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	OA	—	OA	—	—	\$	NYC	71
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	LP	—	L	—	—	\$	CDU, NYC	127
✓	✓	✓	✓	—	✓	C	✓	✓	✓	✓	✓	🚗🚗	OA	—	OA	—	—	🕒, km	CHI	172
✓	✓	✓	✓	—	✓	V	✓	✓	✓	✓	✓	🚗🚗	OA	—	—	—	—	🕒	BE	173, 174
✓	✓	✓	✓	—	✓	C	✓	✓	✓	✓	✓	🚗🚗	LP	—	—	—	—	🕒, km, ☆	NYC	175
✓	✓	✓	✓	—	✓	V	✓	✓	✓	✓	✓	🚗🚗	OA	—	—	—	—	\$	AR	176
✓	✓	✓	✓	—	✓	Q	✓	✓	✓	✓	✓	🚗🚗	NLP	—	—	—	NLP	🕒	AR	96
✓	✓	✓	✓	—	✓	V	✓	✓	✓	✓	✓	🚗🚗	ILP	—	—	—	—	🕒	NYC	77
✓	✓	✓	✓	—	✓	V	✓	✓	✓	✓	✓	🚗🚗	OA	—	—	—	—	🕒, \$	NYC	104
✓	✓	✓	✓	—	✓	V	✓	✓	✓	✓	✓	—	—	L	—	—	L	km, ☆	CO	105
✓	✓	✓	✓	—	✓	V	✓	✓	✓	✓	✓	—	—	L	—	—	—	🕒	AR	116
✓	✓	✓	✓	—	✓	C	✓	✓	✓	✓	✓	—	—	—	—	—	—	☆	SG	117
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	—	—	LP	—	—	🕒	AR	35
✓	✓	✓	✓	—	✓	Q	✓	✓	✓	✓	✓	🚗🚗	—	—	LP	—	—	☆	NYC	82
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	—	—	MINLP	—	—	\$	SG	70
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	—	—	OA	—	—	🕒, ☆	SFO	86
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	—	—	LP	—	—	🚗	AR	118
✓	✓	✓	✓	—	✓	Q	✓	✓	✓	✓	✓	🚗🚗	—	—	LP	—	—	\$	HZ	177
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	—	—	NLP	—	NLP	🕒	HW, CDU	68
✓	✓	✓	✓	—	✓	V	✓	✓	✓	✓	✓	🚗🚗	—	—	ILP	—	—	☆	NYC	111
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	—	—	MINLP	—	—	☆	AR	178
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	—	—	LP	—	—	🕒, ☆	SE	179
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	—	—	LP	—	—	☆	SG	180
✓	✓	✓	✓	—	✓	F	✓	✓	✓	✓	✓	🚗🚗	—	—	—	—	—	\$	AU	43

^a🔌, solutions for problems related to charging of autonomous vehicles (AVs); 🅑, solutions for scheduling problems related to parking of AVs.

^bT, time; D, demand; C, congestion.

^cC, continuum; F, fluidic; Q, queueing-theoretic; V, vehicle-centric.

^d🚗, fleet size; 🅑, parking; 🅑, limited AV autonomy.

^eCO, optimization problem with convex cost and constraints; ILP, optimization problem with integer decision variables and linear cost and constraints; L, learning (e.g., reinforcement learning);

LP, optimization problem with linear cost and constraints; MILP, optimization problem with integer and real decision variables and linear cost and constraints; MINLP, optimization problem with integer and real decision variables and nonlinear cost and constraints; NLP, optimization problem with nonlinear cost and constraints; OA, own algorithms (e.g., heuristic, provable algorithms);

QCILP, optimization problem with integer decision variables, linear cost, and quadratic constraints; QP, optimization problem with quadratic cost and linear constraints.

^f\$, revenue and costs; ⊙, waiting and travel time; km, travel distance; ☆, demand satisfaction; 🚗, fleet size; SW, social welfare.

^gAR, artificial; AU, Austin; BE, Berlin; CA, California; CDU, Chengdu; CHI, Chicago; CO, Cologne; DFW, Dallas-Fort Worth; HZ, Hangzhou; LA, Los Angeles; MIT, Massachusetts

Institute of Technology campus; NB, Newport Beach; NYC, New York City; OC, Orange County; SE, Seattle; SF, Sioux Falls; SFO, San Francisco; SG, Singapore; TO, Tokyo; W, Wuhan.

can be easily and scalably solved as linear programs. To further improve scalability, a variety of algorithms have been implemented (28–31, 104). Recent works have investigated learning-based approaches (e.g., learning optimal policies via neural networks) to improve real-time capabilities (77, 105).

4.1.2. Routing. Routing of AVs is an instance of a dynamic vehicle routing problem, part of the larger family of vehicle routing problems (32). Vehicle routing problems are usually solved (globally or via heuristics) as static routing problems, in which trip origins and destinations are known in advance (for an extensive review, see 106). Crucially, however, in AMoD systems demand is dynamic, leading to dynamic vehicle routing problems. In particular, dynamic vehicle routing problems where people or goods need to be transported between an origin and a destination are dynamic pickup and delivery problems (33), which can be subdivided into three classes: many-to-many problems, in which each request is routed to multiple destinations (107); one-to-one problems, in which each request has a single origin and a single destination (84); and one-to-many problems, in which goods initially located at a depot are delivered to customers, and goods originating from customers are collected and delivered to the depot (108). Clearly, routing AVs is a dynamic one-to-one pickup and delivery problem. However, a full characterization of AMoD systems requires additional properties. First, requests can be immediate and for immediate service, and not meant as prior booking. Second, arrival and travel times are stochastic (16). Finally, apart from ride-sharing cases, AVs are assumed to have single-party occupancy. In AMoD systems, one computes policies (and not routes) to service requests arriving during operations (18, 109). We identify two approaches: custom algorithms (including heuristics and machine learning) (110, 111) and algorithmic queuing theory (18, 112, 113).

Custom algorithms for routing AVs have been studied, for instance, by Jaillet & Wagner (78) and Levin et al. (102). Often, these algorithms are assessed via comparisons with omniscient (noncausal) algorithms or supported by extensive empirical experiments, rarely enjoying theoretic guarantees (102, 114–117). Algorithmic queuing theory usually builds on continuum models of the transportation system, where AVs move at a constant velocity and demand is modeled as a spatiotemporal Poisson process. As optimizing over all policies is intractable, research has focused on designing algorithms, providing rigorous theoretic guarantees on achievable performance, and creating extensions (e.g., to obstacles, time constraints, and priorities) (112, 113). However, most formal performance guarantees for algorithms developed using algorithmic queuing theory hold only in asymptotic regimes, such as a light or heavy load. In medium-load regimes, where systems usually operate, it is difficult to assess the performance of such algorithms. Furthermore, it is hard to incorporate additional operational complexity, such as congestion and charging constraints (81).

4.1.3. Rebalancing. AMoD systems tend to become out of balance: When some origins and destinations are more popular than others, AVs build up at some stations and become depleted at others (35). Historically, researchers studied rebalancing for one-way car-sharing schemes (e.g., car2go and Zipcar) and bike-sharing (107, 111). While these systems also suffer from asymmetric demand, they perform infrequent rebalancing (e.g., happening a few times a day). Thus, the continuous operation of AMoD systems calls for new tools to reason about the relocation of customer-free AVs.

In queuing-theoretic settings, rebalancing strategies are computed via linear programs. The resulting real-time implementation is formally an integer linear program, which, however, reduces to a linear program thanks to total unimodularity (93). In network flow models, rebalancing problems are often formulated as (integer) optimization problems (72, 94, 118), possibly

considering competitive behaviors (119). For instance, recall the dynamic model in Equation 1. Here, the optimal rebalancing strategy α_{ij} from node i to node j results from the linear program

$$\begin{aligned} & \min_{\alpha_{ij} \geq 0} \sum_{i \in S} \sum_{j \in D} t_{ij} \alpha_{ij} \\ & \text{subject to } \sum_{j \in D} \alpha_{ij} = -\lambda_i + \sum_{j \neq i} \lambda_j p_{ji} \quad \forall i \in S := \{i \in \{1, \dots, N\} : \lambda_i < \lambda_1 p_{1i} + \dots + \lambda_N p_{Ni}\}, \\ & \sum_{i \in S} \alpha_{ij} = \lambda_j - \sum_{i \neq j} \lambda_i p_{ij} \quad \forall j \in D := \{1, \dots, N\} \setminus S. \end{aligned}$$

In most cases, real-time rebalancing has been studied in a model predictive control framework, sometimes combined with stochastic models for future demand, possibly enhanced by machine learning (120, 121). Previous works have also exploited total unimodularity to reduce integer linear programs to linear programs (122).

4.1.4. Dispatching, routing, and rebalancing. Most works have focused on solving routing, dispatching, and rebalancing problems together, sometimes along with ride-sharing problems, as well as on incorporating specific operational constraints. For instance, in network flow models, we lose the identity of single AVs. Thus, the problems of dispatching, routing, and rebalancing AVs are implicitly solved when computing minimum cost flows, which are usually formulated as linear programs and quadratic programs for noninteger flows (72, 94, 122) and as integer linear programs and mixed-integer linear programs for integer flows (120), possibly considering charging (123, 124) and parking routines (69).

In time-invariant settings, Salazar et al. (56) and Wollenstein-Betech et al. (57) proposed a quadratic program for socially optimal dispatching, routing, and rebalancing, considering interactions with other modes of transportation (e.g., public transport). Among other things, the authors outlined the benefits of cooperatively operating AMoD systems and public transport in New York City and Berlin. Similarly, Rossi et al. (55) and Salazar et al. (92) solved the problems via linear programs, and Solovey et al. (125) did so via convex optimization. Interestingly, the time-invariant setting enables one to account for endogenous congestion effects. With the exception of the paper by Rossi et al. (55), which proposed a real-time implementation based on an integer linear program, these works are meant not to devise operational policies but to enable planning studies.

In time-varying settings, the problems are often cast within a model predictive control framework: Carron et al. (94) formulated a quadratic program, and Kang & Levin (69), Zraggen et al. (72), and Tsao et al. (122) formulated a linear program, sometimes accounting for parking. Similarly, Boewing et al. (76), Zhang et al. (123), and Iacobucci et al. (124) framed the problem as a mixed-integer linear program, considering the charging of electric AVs. Ma et al. (101) proposed a linear program, and Yang et al. (95) proposed a model predictive control framework based on Stackelberg games, solved via a linear program. Miao et al. (126) adopted a distributionally robust approach with respect to uncertainty in future demand, solving the problem via convex optimization, and Gu eriau et al. (115) offered a learning-based solution that uses neural networks, together with microscopic simulations of the transportation system, to learn optimal policies. All of these approaches are meant for real-time implementation, possibly up to relaxations, heuristics, or suboptimality of the underlying optimization problems. Heuristic algorithms were proposed by Hyland & Mahmassani (79). Finally, Gammelli et al. (127) decoupled dispatching and rebalancing (neglecting routing) in two linear programs and linked them through reinforcement learning: Once AVs are dispatched by solving the first linear program, the optimal

distribution of customer-free AVs is computed via reinforcement learning (based on graph neural networks) and realized by solving the second linear program.

Other optimization-based approaches have been proposed for planning studies (128, 129), possibly in combination with the power grid (53, 54). In these studies, for instance, the authors showed that for a case study in Orange County, California, coordination between AVs and the power grid eliminated 99% of overloads and 50% of voltage drops.

4.1.5. Dispatching, routing, and rebalancing with ride-sharing. Tsao et al. (73) framed dispatching, routing, rebalancing, and ride-sharing as network flow problems and formulated a mixed-integer linear program to optimize the costs and times of operations, implemented in real time in a receding-horizon fashion that relies on forecasting for the demand. By contrast, Miller & How (130) relied on integer programming and, to overcome the computational burden, heuristics. Alonso-Mora et al. (38), Sayarshad & Chow (110), and Ma et al. (131) broke the problem down into two instances. Alonso-Mora et al. (38) solved dispatching, routing, and ride-sharing via an integer linear program and solved rebalancing through a linear program. Among other things, the authors showed that, without ride-sharing, 98% of the taxi rides in New York City, currently served by more than 13,000 taxis, could be served with just 3,000 AVs, and that ride-sharing leads to substantial additional benefits (e.g., less travel distance). Fielbaum et al. (132) extended the work of Alonso-Mora et al. (38) to assess the benefits of including walking sections to maximize system performance. Sayarshad & Chow (110) and Ma et al. (131) solved dispatching, routing, and ride-sharing via an ad hoc heuristic algorithm and solved rebalancing via a mixed-integer linear program, whose solutions are computed approximately via Lagrange decomposition. Levin et al. (102) and Liu et al. (103) relied entirely on ad hoc algorithms, and Lin et al. (114) relied on reinforcement learning. Finally, Chu et al. (74) combined graph-theoretic and vehicle-centric models to formulate the problem as an integer linear program, including the dynamic state of charge of the electric AVs and parking.

4.2. Existing Tools for Simulation

There are two types of simulation tools for AMoD systems: custom-made and general transportation simulators. While custom-made simulators include ad hoc software implementations to numerically analyze and control AMoD systems, they are typically not adaptable to novel scenarios, they do not generalize, and, in some cases, they are simplistic. General transportation simulators, by contrast, entail greater software complexity and result from years of development and validation (133–135). Simulators can be further divided into three classes: microscopic, macroscopic, and mesoscopic. Microscopic simulators model both AVs and traffic dynamics (136) and are commonly used when modeling microscopic effects such as navigation of intersections or merging on freeways (135, 137–139). Macroscopic simulators are designed to analyze effects at a higher abstraction level (140); transportation systems are generally abstracted as coarse-grained graphs, and agents are represented as flows. Mesoscopic simulators operate at an intermediate abstraction level, omitting the dynamics of individual AVs but maintaining a faithful representation of the network, demand, and fleet size (45, 56, 141, 142).

Researchers have identified three ways to validate algorithms: synthetic, hardware, and real-world case studies. Synthetic case studies rely on the generation of artificial networks and demand and allow for quick prototyping and proofs of concept (e.g., scalability). Hardware tests provide more practical observations but are feasible only on small scales (82, 143, 144). For these reasons, most works have presented real-world case studies, usually constructed as follows. Road networks can be downloaded from OpenStreetMap (145), and schedules of public transportation systems

are often publicly available as General Transit Feed Specification (GTFS) data (146). Hence, the availability of demand data drives the design of the case study, and, in particular, the choice of city.³ New York City has been a pioneer, releasing taxi demand data via the New York City Taxi and Limousine Commission every year since 2009. The data sets include 697,622,444 trips (147). Thus, the vast majority of the studies in **Table 1** are based in New York City, but more recently, case studies have also been performed for numerous other cities.

5. FUTURE RESEARCH AVENUES

We identify four main avenues for future research: modeling, co-design of AV-enabled transportation systems, interactions among stakeholders in the mobility ecosystem, and fairness, privacy, and trust.

5.1. Modeling

The need for more efficient analysis and control of AMoD systems calls for richer models in terms of demand, congestion, operational constraints, and ride-sharing.

5.1.1. Elastic demand. So far, researchers have considered mainly inelastic demand. The study of elastic demand for AMoD systems is therefore a natural direction for future research. This is an interdisciplinary research area, touching upon topics such as incentives, customer preferences, and mode choice (51, 63, 99, 148, 149). Questions of interest include how AMoD systems can attract new customers, what effective tools a mobility provider can use to steer mode choice, and whether AMoD systems are scalable to accommodate induced demand.

5.1.2. Congestion. As suggested by **Table 1**, most existing operational policies ignore endogenous congestion. While in the last few years more research in this direction has been pursued [e.g., in the seminal work by Rossi (150)], more questions need to be answered, mainly along two directions: (a) approaches to reduce congestion, including ride-sharing, demand staggering, and integration with other mobility systems (e.g., public transport and micromobility), and (b) computationally efficient methods for more accurate congestion models (125).

5.1.3. Operational constraints. While considerable attention has been dedicated to fleet size and charging constraints, other important aspects have been ignored. We identify two main research gaps.

5.1.3.1. Autonomy-related limits. First, high usage of AVs requires not only scheduling recharging or refueling sessions but also considering autonomy-related operational limits, such as energy consumption resulting from intense sensing and computing [autonomy is estimated to increase a vehicle's energy consumption by 20% (151)], and (autonomous) calibration (58, 152). Second, one needs to strategically locate (autonomous) calibration facilities and efficiently schedule calibration sessions. Facilities could also be used for general AV maintenance and cleaning. Indeed, in an AV, no driver can ensure that customers leave the vehicle clean and ready for the next customers.

³We have created a list of existing publicly available demand data sets for the simulation of AMoD systems; see <http://gioele.science/amod-demand-data>.

5.1.3.2. Autonomy-enabling infrastructure. In this article, we have tacitly assumed that automating vehicles is the only feasible way to achieve AMoD. Yet this is just part of the story. One interesting alternative is autonomy-enabling infrastructure, in which part of the autonomy pipeline of AVs is outsourced to the public infrastructure [e.g., online, cloud-based perception and mapping (59, 153)]. Open questions include whether this approach is feasible, how much it would cost, and whether it would scale effectively.

5.1.4. Ride-sharing. Various studies have outlined potential benefits and challenges of ride-sharing (36–38, 154). While recent studies are a first step in the systematic assessment of ride-sharing (36), further work needs to be done to determine to what extent ride-sharing is beneficial and whether it is reasonable only for certain urban environments (e.g., only for specific network topologies).

5.2. Co-Design of Autonomous Vehicle-Enabled Transportation Systems

The current design process for future mobility solutions lacks clear, specific requirements for the service they would be providing (14). Indeed, knowledge about the service could drastically impact their design. For instance, knowing that an AV does not need to reach speeds higher than 30 mph (e.g., due to slow average speeds in cities) might significantly simplify its design and accelerate its deployment (39, 45). Furthermore, from the system-level perspective of transportation management, knowledge about the development of new mobility approaches has an impact on future infrastructure investments and provisions of service. Future mobility solutions need to be co-designed with the transportation system they enable. For the particular case of AMoD systems, this requires tools to study their design all the way from the autonomy of AVs to their effects on fleet control and on the transportation system at large (46, 155). Existing approaches for such coupled problems typically have a fixed, problem-specific structure and therefore do not allow one to rigorously design the transportation infrastructure in a modular and compositional fashion (45, 156, 157).

5.3. Interactions Among Stakeholders in the Mobility Ecosystem

The rise of private, profit-maximizing mobility providers leveraging public resources, such as ride-hailing companies and, in the near future, AMoD systems, raises questions for current regulation schemes, such as whether mobility providers should be taxed for usage of public resources and how many AVs a city should allow. These questions call for tools to reason about interactions among stakeholders in the mobility ecosystem, including municipalities, mobility providers, and customers. First, municipalities need to design regulatory policies tailored to new mobility options, such as limitations on fleet sizes, public transport prices, and taxes (51, 63–65, 158, 159). Second, different mobility providers will operate in the same environment, opening the field for competitive analyses (64, 65, 160, 161). Third, new mobility approaches will influence travel needs. Finally, coordination schemes for AMoD systems could relate to transportation of goods. Examples include delivery of packages (47) and shipping (162).

5.4. Fairness, Privacy, and Trust

The penetration of AMoD systems calls for considerations of fairness, privacy, and trust. First, fairness is an important aspect for centrally controlled systems; for example, the discrepancy among users' travel times can reach 100% in real-world transportation networks (163).

Importantly, stakeholders need to find the right trade-off between fairness and efficiency in the transportation system and study incentive mechanisms to steer mobility providers accordingly (164, 165). Furthermore, while customers' data can be used to improve the service quality of transportation systems, privacy and trust concerns might arise (166–168). Privacy includes handling methodologies for data containing sensitive information, and trust concerns appear as soon as mobility providers and regulation authorities have misaligned interests (167, 169).

6. CONCLUSION

In this review, we have provided a full picture of the analysis and control of AMoD systems by identifying relevant problem settings, giving an overview of mathematical models, and analyzing existing solution approaches, highlighting results, trends, and limitations. Finally, we have armed the reader with insights on future research.

SUMMARY POINTS

- **Problems:** The literature on the analysis and control of autonomous mobility-on-demand (AMoD) systems focuses on two classes of problems: operational policies [e.g., how to route autonomous vehicles (AVs)] and strategic planning (e.g., where to place charging stations).
- **Mathematical models:** Modeling AMoD systems requires models of the transportation network, AVs, demand, congestion, operational constraints, and interactions with existing infrastructure.
- **Solution approaches:** Researchers have developed a wide variety of methods to solve problems related to AMoD systems, relying on mathematical optimization, custom algorithms, and machine learning. Performance metrics include costs, waiting and travel time, travel distance, demand satisfaction, fleet size, and social welfare.

FUTURE ISSUES

- **Models:** Effects such as elastic demand, endogenous congestion, operational constraints, and ride-sharing require further investigation.
- **Co-design:** Mobility solutions need to be co-designed with the system they enable.
- **Interactions:** We need tools to rigorously reason about interactions among stakeholders in the mobility ecosystem.
- **Fairness, privacy, and trust:** Deployment of AMoD raises concerns regarding fairness among customers as well as privacy and trust when handling private data.

DISCLOSURE STATEMENT

E.F. was chief scientist for Motional and its predecessors (Aptiv Autonomous Driving, nuTonomy) from August 2013 through March 2021. M.P. is the director of the Autonomous Vehicle Research group at Nvidia Corporation.

ACKNOWLEDGMENTS

The authors' research is supported by the Swiss National Science Foundation under National Centre of Competence in Research (NCCR) Automation (grant agreement 51NF40_180545) and the National Science Foundation [Faculty Early Career Development Program (CAREER) award CMMI1454737]. The authors thank Sonia Monti for graphics support and Devansh Jalota for useful discussions.

LITERATURE CITED

1. UN Dep. Econ. Soc. Aff. 2018. 68% of the world population projected to live in urban areas by 2050, says UN. *United Nations Department of Economic and Social Affairs*, May 16. <https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html>
2. UN. 2021. Cities and pollution. *United Nations*. <https://www.un.org/en/climatechange/climate-solutions/cities-pollution>
3. Schrank D, Eisele B, Lomax T. 2019. *Urban mobility report 2019*. Rep., Tex. A&M Transp. Inst., College Station
4. Mitchell WJ, Borroni-Bird CE, Burns LD. 2010. *Reinventing The Automobile: Personal Urban Mobility for the 21st Century*. Cambridge, MA: MIT Press
5. Pavone M. 2015. Autonomous mobility-on-demand systems for future urban mobility. In *Autonomes Fahren: technische, rechtliche und gesellschaftliche Aspekte*, ed. M Maurer, J Gerdes, B Lenz, H Winner, pp. 399–416. Berlin: Springer
6. City of New York. 2021. *OneNYC 2050: building a strong and fair city*. Rep., City of New York. <http://onenyc.cityofnewyork.us/reports-resources/>
7. Berger T, Chen C, Frey CB. 2018. Drivers of disruption? Estimating the Uber effect. *Eur. Econ. Rev.* 110:197–210
8. Rogers B. 2015. The social costs of Uber. *Univ. Chicago Law Rev. Online* 82:85–102
9. AAA Found. Traffic Saf. 2021. *New American driving survey: updated methodology and results from July 2019 to June 2020*. Tech. Rep., AAA Found. Traffic Saf., Washington, DC. <https://aaafoundation.org/new-american-driving-survey-updated-methodology-and-results-from-july-2019-to-june-2020>
10. Hancock PA, Nourbakhsh I, Stewart J. 2019. On the future of transportation in an era of automated and autonomous vehicles. *PNAS* 116:7684–91
11. Becker H, Becker F, Abe R, Bekhor S, Belgiawan PF, et al. 2020. Impact of vehicle automation and electric propulsion on production costs for mobility services worldwide. *Transp. Res. A* 138:105–26
12. Iagnemma K. 2020. 100,000 self-driving rides strong. *Aptiv*, Feb. 11. <https://www.apativ.com/en/insights/article/100000-Self-Driving-Rides-Strong>
13. Schwall M, Daniel T, Victor T, Favaro F, Hohnhold H. 2020. Waymo public road safety performance data. arXiv:2011.00038 [cs.RO]
14. Yigitcanlar T, Wilson M, Kamruzzaman M. 2019. Disruptive impacts of automated driving systems on the built environment and land use: an urban planner's perspective. *J. Open Innov. Technol. Mark. Complex.* 5:24
15. Kim G, Ong YS, Heng CK, Tan PS, Zhang NA. 2015. City vehicle routing problem (city VRP): a review. *IEEE Trans. Intell. Transp. Syst.* 16:1654–66
16. Psaraftis HN, Wen M, Kontovas CA. 2016. Dynamic vehicle routing problems: three decades and counting. *Networks* 67:3–31
17. Eksiöglu B, Vural AV, Reisman A. 2009. The vehicle routing problem: a taxonomic review. *Comput. Ind. Eng.* 57:1472–83
18. Bullo F, Frazzoli E, Pavone M, Savla K, Smith SL. 2011. Dynamic vehicle routing for robotic systems. *Proc. IEEE* 99:1482–504
19. Gendreau M, Potvin JY. 1998. Dynamic vehicle routing and dispatching. In *Fleet Management and Logistics*, ed. TG Crainic, G Laporte, pp. 115–26. Berlin: Springer

20. Schiffer M, Schneider M, Walther G, Laporte G. 2019. Vehicle routing and location routing with intermediate stops: a review. *Transp. Sci.* 53:319–43
21. Zhao L, Malikopoulos A. 2020. Enhanced mobility with connectivity and automation: a review of shared autonomous vehicle systems. *IEEE Intell. Transp. Syst. Mag.* 14(1):87–102
22. Narayanan S, Chaniotakis E, Antoniou C. 2020. Shared autonomous vehicle services: a comprehensive review. *Transp. Res. C* 111:255–93
23. Hyland MF, Mahmassani HS. 2017. Taxonomy of shared autonomous vehicle fleet management problems to inform future transportation mobility. *Transp. Res. Rec.* 2653:26–34
24. Mourad A, Puchinger J, Chu C. 2019. A survey of models and algorithms for optimizing shared mobility. *Transp. Res. B* 123:323–46
25. Golbabaei F, Yigitcanlar T, Bunker J. 2021. The role of shared autonomous vehicle systems in delivering smart urban mobility: a systematic review of the literature. *Int. J. Sustain. Transp.* 15:731–48
26. Markov I, Guglielmetti R, Laumanns M, Fernández-Antolín A, de Souza R. 2021. Simulation-based design and analysis of on-demand mobility services. *Transp. Res. A* 149:170–205
27. Agatz N, Erera A, Savelsbergh M, Wang X. 2012. Optimization for dynamic ride-sharing: a review. *Eur. J. Oper. Res.* 223:295–303
28. Motzkin T. 1956. The assignment problem. *Proc. Symp. Appl. Math.* 6:109–25
29. Kuhn HW. 1955. The Hungarian method for the assignment problem. *Naval Res. Logist. Q.* 2:83–97
30. Bertsekas DP, Castañón DA. 1993. Parallel asynchronous Hungarian methods for the assignment problem. *ORSA J. Comput.* 5:261–74
31. Peeta S, Ziliaskopoulos AK. 2001. Foundations of dynamic traffic assignment: the past, the present and the future. *Netw. Spat. Econ.* 1:233–65
32. Toth P, Vigo D. 2014. *Vehicle Routing: Problems, Methods, and Applications*. Philadelphia: Soc. Ind. Appl. Math.
33. Berbeglia G, Cordeau JF, Laporte G. 2010. Dynamic pickup and delivery problems. *Eur. J. Oper. Res.* 202:8–15
34. Laporte G. 2009. Fifty years of vehicle routing. *Transp. Sci.* 43:408–16
35. Pavone M, Smith SL, Frazzoli E, Rus D. 2012. Robotic load balancing for mobility-on-demand systems. *Int. J. Robot. Res.* 31:839–54
36. Ruch C, Lu C, Sieber L, Frazzoli E. 2021. Quantifying the efficiency of ride sharing. *IEEE Trans. Intell. Transp. Syst.* 22:5811–16
37. Santi P, Resta G, Szell M, Sobolevsky S, Strogatz SH, Ratti C. 2014. Quantifying the benefits of vehicle pooling with shareability networks. *PNAS* 111:13290–94
38. Alonso-Mora J, Samaranayake S, Wallar A, Frazzoli E, Rus D. 2017. On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *PNAS* 114:462–67
39. Spieser K, Treleaven K, Zhang R, Frazzoli E, Morton D, Pavone M. 2014. Toward a systematic approach to the design and evaluation of automated mobility-on-demand systems: a case study in Singapore. In *Road Vehicle Automation*, ed. G Meyer, S Beiker, pp. 229–45. Cham, Switz.: Springer
40. Vazifteh MM, Santi P, Resta G, Strogatz SH, Ratti C. 2018. Addressing the minimum fleet problem in on-demand urban mobility. *Nature* 557:534–38
41. Barrios JA, Godier JD. 2014. Fleet sizing for flexible carsharing systems: simulation-based approach. *Transp. Res. Rec.* 2416:1–9
42. Boesch PM, Ciari F, Axhausen KW. 2016. Autonomous vehicle fleet sizes required to serve different levels of demand. *Transp. Res. Rec.* 2542:111–19
43. Fagnant DJ, Kockelman KM. 2018. Dynamic ride-sharing and fleet sizing for a system of shared autonomous vehicles in Austin, Texas. *Transportation* 45:143–58
44. Wallar A, Schwarting W, Alonso-Mora J, Rus D. 2019. Optimizing multi-class fleet compositions for shared mobility-as-a-service. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 2998–3005. Piscataway, NJ: IEEE
45. Zardini G, Lanzetti N, Censi A, Frazzoli E, Pavone M. 2020. Co-design to enable user-friendly tools to assess the impact of future mobility solutions. arXiv:2008.08975 [eess.SY]

46. Zardini G, Milojevic D, Censi A, Frazzoli E. 2021. Co-design of embodied intelligence: a structured approach. arXiv:2011.10756 [cs.RO]
47. Choudhury S, Solovey K, Kochenderfer MJ, Pavone M. 2021. Efficient large-scale multi-drone delivery using transit networks. *J. Artif. Intell. Res.* 70:757–88
48. Ruch C, Ehrler R, Hörl S, Balac M, Frazzoli E. 2021. Simulation-based assessment of parking constraints for automated mobility on demand: a case study of Zurich. *Vehicles* 3:272–86
49. Zhang W, Guhathakurta S. 2017. Parking spaces in the age of shared autonomous vehicles: How much parking will we need and where? *Transp. Res. Rec.* 2651:80–91
50. Zhang H, Sheppard CJ, Lipman TE, Moura SJ. 2019. Joint fleet sizing and charging system planning for autonomous electric vehicles. *IEEE Trans. Intell. Transp. Syst.* 21:4725–38
51. Chen TD, Kockelman KM. 2016. Management of a shared autonomous electric vehicle fleet: implications of pricing schemes. *Transp. Res. Rec.* 2572:37–46
52. Chen TD, Kockelman KM, Hanna JP. 2016. Operations of a shared, autonomous, electric vehicle fleet: implications of vehicle & charging infrastructure decisions. *Transp. Res. A* 94:243–54
53. Rossi F, Iglesias R, Alizadeh M, Pavone M. 2019. On the interaction between autonomous mobility-on-demand systems and the power network: models and coordination algorithms. *IEEE Trans. Control Netw. Syst.* 7:384–97
54. Estandia A, Schiffer M, Rossi F, Luke J, Kara EC, et al. 2021. On the interaction between autonomous mobility on demand systems and power distribution networks—an optimal power flow approach. *IEEE Trans. Control Netw. Syst.* 8:1163–76
55. Rossi F, Zhang R, Hindy Y, Pavone M. 2018. Routing autonomous vehicles in congested transportation networks: structural properties and coordination algorithms. *Auton. Robots* 42:1427–42
56. Salazar M, Lanzetti N, Rossi F, Schiffer M, Pavone M. 2019. Intermodal autonomous mobility-on-demand. *IEEE Trans. Intell. Transp. Syst.* 21:3946–60
57. Wollenstein-Betech S, Salazar M, Houshmand A, Pavone M, Paschalidis IC, Cassandras CG. 2021. Routing and rebalancing intermodal autonomous mobility-on-demand systems in mixed traffic. *IEEE Trans. Intell. Transp. Syst.* In press. <https://doi.org/10.1109/TITS.2021.3112106>
58. Levinson J, Thrun S. 2013. Automatic online calibration of cameras and lasers. In *Robotics: Science and Systems IX*, ed. P Newman, D Fox, D Hsu, pap. 29. N.p.: Robot. Sci. Syst. Found.
59. Coutinho RW, Boukerche A. 2019. Guidelines for the design of vehicular cloud infrastructures for connected autonomous vehicles. *IEEE Wirel. Commun.* 26:6–11
60. Basu R, Araldo A, Akkinipally AP, Nahmias Biran BH, Basak K, et al. 2018. Automated mobility-on-demand versus mass transit: a multi-modal activity-driven agent-based simulation approach. *Transp. Res. Rec.* 2672:608–18
61. Wen J, Chen YX, Nassir N, Zhao J. 2018. Transit-oriented autonomous vehicle operation with integrated demand-supply interaction. *Transp. Res. C* 97:216–34
62. Alizadeh M, Wai HT, Chowdhury M, Goldsmith A, Scaglione A, Javidi T. 2016. Optimal pricing to manage electric vehicles in coupled power and transportation networks. *IEEE Trans. Control Netw. Syst.* 4:863–75
63. Turan B, Pedarsani R, Alizadeh M. 2020. Dynamic pricing and fleet management for electric autonomous mobility on demand systems. *Transp. Res. C* 121:102829
64. Lanzetti N, Schiffer M, Ostrovsky M, Pavone M. 2020. *On the interplay between self-driving cars and public transportation: a game theoretic perspective*. Work. Pap., Cah. GERAD G-2020-24, HEC Montréal, Montreal, Can. <https://www.gerad.ca/en/papers/G-2020-24>
65. Zardini G, Lanzetti N, Guerrini L, Frazzoli E, Dörfler F. 2021. Game theory to study interactions between mobility stakeholders. In *2021 IEEE 24th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2054–61. Piscataway, NJ: IEEE
66. Neuburger H. 1971. The economics of heavily congested roads. *Transp. Res.* 5:283–93
67. Ford LR Jr. 1956. *Network flow theory*. Tech. Rep. P-923, RAND Corp., Santa Monica, CA
68. Guo G, Xu T. 2020. Vehicle rebalancing with charging scheduling in one-way car-sharing systems. *IEEE Trans. Intell. Transp. Syst.* In press. <https://doi.org/10.1109/TITS.2020.3043594>
69. Kang D, Levin MW. 2021. Maximum-stability dispatch policy for shared autonomous vehicles. *Transp. Res. B* 148:132–51

70. Nair R, Miller-Hooks E. 2011. Fleet management for vehicle sharing operations. *Transp. Sci.* 45:524–40
71. Duan L, Wei Y, Zhang J, Xia Y. 2020. Centralized and decentralized autonomous dispatching strategy for dynamic autonomous taxi operation in hybrid request mode. *Transp. Res. C* 111:397–420
72. Zraggen J, Tsao M, Salazar M, Schiffer M, Pavone M. 2019. A model predictive control scheme for intermodal autonomous mobility-on-demand. In *2019 22nd IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 1953–60. Piscataway, NJ: IEEE
73. Tsao M, Mилоjevic D, Ruch C, Salazar M, Frazzoli E, Pavone M. 2019. Model predictive control of ride-sharing autonomous mobility-on-demand systems. In *2019 International Conference on Robotics and Automation (ICRA)*, pp. 6665–71. Piscataway, NJ: IEEE
74. Chu KF, Lam AY, Li VO. 2021. Joint rebalancing and vehicle-to-grid coordination for autonomous vehicle public transportation system. *IEEE Trans. Intell. Transp. Syst.* In press. <https://doi.org/10.1109/TITS.2021.3067044>
75. James J, Lam AY. 2017. Autonomous vehicle logistic system: joint routing and charging strategy. *IEEE Trans. Intell. Transp. Syst.* 19:2175–87
76. Boewing F, Schiffer M, Salazar M, Pavone M. 2020. A vehicle coordination and charge scheduling algorithm for electric autonomous mobility-on-demand systems. In *2020 American Control Conference (ACC)*, pp. 248–55. Piscataway, NJ: IEEE
77. Hu L, Dong J. 2022. An artificial-neural-network-based model for real-time dispatching of electric autonomous taxis. *IEEE Trans. Intell. Transp. Syst.* 23:1519–28
78. Jaillet P, Wagner MR. 2006. Online routing problems: value of advanced information as improved competitive ratios. *Transp. Sci.* 40:200–10
79. Hyland M, Mahmassani HS. 2018. Dynamic autonomous vehicle fleet operations: optimization-based strategies to assign AVs to immediate traveler demand requests. *Transp. Res. C* 92:278–97
80. George DK. 2012. *Stochastic modeling and decentralized control policies for large-scale vehicle sharing systems via closed queueing networks*. PhD Thesis, Ohio State Univ., Columbus
81. Zhang R. 2016. *Models and large-scale coordination algorithms for autonomous mobility-on-demand*. PhD Thesis, Stanford Univ., Stanford, CA
82. Zhang R, Pavone M. 2016. Control of robotic mobility-on-demand systems: a queueing-theoretical perspective. *Int. J. Robot. Res.* 35:186–203
83. Ruch C, Gächter J, Hakenberg J, Frazzoli E. 2020. The +1 method: model-free adaptive repositioning policies for robotic multi-agent systems. *IEEE Trans. Netw. Sci. Eng.* 7:3171–84
84. Treleven K, Pavone M, Frazzoli E. 2013. Asymptotically optimal algorithms for one-to-one pickup and delivery problems with applications to transportation systems. *IEEE Trans. Autom. Control* 58:2261–76
85. Levina E, Bickel P. 2001. The Earth Mover's distance is the Mallows distance: some insights from statistics. In *2001 8th IEEE International Conference on Computer Vision (ICCV)*, Vol. 2, pp. 251–56. Piscataway, NJ: IEEE
86. Albert M, Ruch C, Frazzoli E. 2019. Imbalance in mobility-on-demand systems: a stochastic model and distributed control approach. *ACM Trans. Spat. Algorithms Syst.* 5:1–22
87. Babonneau F, Vial JP. 2008. An efficient method to compute traffic assignment problems with elastic demands. *Transp. Sci.* 42:249–60
88. Dafermos SC, Sparrow FT. 1969. The traffic assignment problem for a general network. *J. Res. Natl. Bur. Stand. B* 73:91–118
89. Gartner NH. 1980. Optimal traffic assignment with elastic demands: a review part I. Analysis framework. *Transp. Sci.* 14:174–91
90. Gartner NH. 1980. Optimal traffic assignment with elastic demands: a review part II. Algorithmic approaches. *Transp. Sci.* 14:192–208
91. US Bur. Public Roads. 1964. *Traffic Assignment Manual for Application with a Large, High Speed Computer*, Vol. 2. Washington, DC: US Gov. Print. Off.
92. Salazar M, Tsao M, Aguiar I, Schiffer M, Pavone M. 2019. A congestion-aware routing scheme for autonomous mobility-on-demand systems. In *2019 18th European Control Conference (ECC)*, pp. 3040–46. Piscataway, NJ: IEEE
93. Zhang R, Rossi F, Pavone M. 2018. Analysis, control, and evaluation of mobility-on-demand systems: a queueing-theoretical approach. *IEEE Trans. Control Netw. Syst.* 6:115–26

94. Carron A, Seccamonte F, Ruch C, Frazzoli E, Zeilinger MN. 2021. Scalable model predictive control for autonomous mobility-on-demand systems. *IEEE Trans. Control Syst. Technol.* 29:635–44
95. Yang K, Tsao MW, Xu X, Pavone M. 2020. Planning and operations of mixed fleets in mobility-on-demand systems. arXiv:2008.08131 [eess.SY]
96. Belakaria S, Ammous M, Smith L, Sorour S, Abdel-Rahim A. 2019. Multi-class management with sub-class service for autonomous electric mobility on-demand systems. *IEEE Trans. Veh. Technol.* 68:7155–59
97. Tucker N, Turan B, Alizadeh M. 2019. Online charge scheduling for electric vehicles in autonomous mobility on demand fleets. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 226–31. Piscataway, NJ: IEEE
98. Cenedese C, Fabiani F, Cucuzzella M, Scherpen JM, Cao M, Grammatico S. 2019. Charging plug-in electric vehicles as a mixed-integer aggregative game. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pp. 4904–9. Piscataway, NJ: IEEE
99. Fagnant DJ, Kockelman KM. 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transp. Res. A* 77:167–81
100. Sieber L, Ruch C, Hörl S, Axhausen KW, Frazzoli E. 2020. Improved public transportation in rural areas with self-driving cars: a study on the operation of Swiss train lines. *Transp. Res. A* 134:35–51
101. Ma J, Li X, Zhou F, Hao W. 2017. Designing optimal autonomous vehicle sharing and reservation systems: a linear programming approach. *Transp. Res. C* 84:124–41
102. Levin MW, Kockelman KM, Boyles SD, Li T. 2017. A general framework for modeling shared autonomous vehicles with dynamic network-loading and dynamic ride-sharing application. *Comput. Environ. Urban Syst.* 64:373–83
103. Liu Z, Miwa T, Zeng W, Bell MG, Morikawa T. 2019. Dynamic shared autonomous taxi system considering on-time arrival reliability. *Transp. Res. C* 103:281–97
104. Li L, Pantelidis T, Chow JY, Jabari SE. 2021. A real-time dispatching strategy for shared automated electric vehicles with performance guarantees. *Transp. Res. E* 152:102392
105. James J, Yu W, Gu J. 2019. Online vehicle routing with neural combinatorial optimization and deep reinforcement learning. *IEEE Trans. Intell. Transp. Syst.* 20:3806–17
106. Cordeau JF, Laporte G, Potvin JY, Savelsbergh MW. 2007. Transportation on demand. *Handb. Oper. Res. Manag. Sci.* 14:429–66
107. Dell’Amico M, Hadjicostantinou E, Iori M, Novellani S. 2014. The bike sharing rebalancing problem: mathematical formulations and benchmark instances. *Omega* 45:7–19
108. Montané FAT, Galvao RD. 2006. A tabu search algorithm for the vehicle routing problem with simultaneous pick-up and delivery service. *Comput. Oper. Res.* 33:595–619
109. Smith SL. 2009. *Task allocation and vehicle routing in dynamic environments*. PhD Thesis, Univ. Calif., Santa Barbara
110. Sayarshad HR, Chow JY. 2017. Non-myopic relocation of idle mobility-on-demand vehicles as a dynamic location-allocation-queueing problem. *Transp. Res. E* 106:60–77
111. Wallar A, Van Der Zee M, Alonso-Mora J, Rus D. 2018. Vehicle rebalancing for mobility-on-demand systems with ride-sharing. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4539–46. Piscataway, NJ: IEEE
112. Pavone M, Bisnik N, Frazzoli E, Isler V. 2009. A stochastic and dynamic vehicle routing problem with time windows and customer impatience. *Mobile Netw. Appl.* 14:350–64
113. Pavone M. 2010. *Dynamic vehicle routing for robotic networks*. PhD Thesis, Mass. Inst. Technol., Cambridge
114. Lin K, Zhao R, Xu Z, Zhou J. 2018. Efficient large-scale fleet management via multi-agent deep reinforcement learning. In *KDD ’18: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1774–83. New York: ACM
115. Guériau M, Cugurullo F, Acheampong RA, Dusparic I. 2020. Shared autonomous mobility on demand: a learning-based approach and its performance in the presence of traffic congestion. *IEEE Intell. Transp. Syst. Mag.* 12(4):208–18
116. Mao C, Shen Z. 2018. A reinforcement learning framework for the adaptive routing problem in stochastic time-dependent network. *Transp. Res. C* 93:179–97

117. Han M, Senellart P, Bressan S, Wu H. 2016. Routing an autonomous taxi with reinforcement learning. In *CIKM '16: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pp. 2421–24. New York: ACM
118. Smith SL, Pavone M, Schwager M, Frazzoli E, Rus D. 2013. Rebalancing the rebalancers: optimally routing vehicles and drivers in mobility-on-demand systems. In *2013 American Control Conference*, pp. 2362–67. Piscataway, NJ: IEEE
119. Martin L. 2020. *Rebalancing in shared mobility systems*. PhD Thesis, Tech. Univ. München, Munich
120. Iglesias R, Rossi F, Wang K, Hallac D, Leskovec J, Pavone M. 2018. Data-driven model predictive control of autonomous mobility-on-demand systems. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6019–25. Piscataway, NJ: IEEE
121. Dandl F, Hyland M, Bogenberger K, Mahmassani HS. 2019. Evaluating the impact of spatio-temporal demand forecast aggregation on the operational performance of shared autonomous mobility fleets. *Transportation* 46:1975–96
122. Tsao M, Iglesias R, Pavone M. 2018. Stochastic model predictive control for autonomous mobility on demand. In *2018 21st IEEE International Conference on Intelligent Transportation Systems (ITSC)*, pp. 3941–48. Piscataway, NJ: IEEE
123. Zhang R, Rossi F, Pavone M. 2016. Model predictive control of autonomous mobility-on-demand systems. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1382–89. Piscataway, NJ: IEEE
124. Iacobucci R, McLellan B, Tezuka T. 2019. Optimization of shared autonomous electric vehicles operations with charge scheduling and vehicle-to-grid. *Transp. Res. C* 100:34–52
125. Solovey K, Salazar M, Pavone M. 2019. Scalable and congestion-aware routing for autonomous mobility-on-demand via Frank-Wolfe optimization. In *Robotics: Science and Systems XV*, ed. A Bicchi, H Kress-Gazit, S Hutchinson, pap. 66. N.p.: Robot. Sci. Syst. Found.
126. Miao F, Han S, Hendawi AM, Khalefa ME, Stankovic JA, Pappas GJ. 2017. Data-driven distributionally robust vehicle balancing using dynamic region partitions. In *2017 ACM/IEEE 8th International Conference on Cyber-Physical Systems*, pp. 261–72. Piscataway, NJ: IEEE
127. Gammelli D, Yang K, Harrison J, Rodrigues F, Pereira FC, Pavone M. 2021. Graph neural network reinforcement learning for autonomous mobility-on-demand systems. arXiv:2104.11434 [eess.SY]
128. Levin MW. 2017. Congestion-aware system optimal route choice for shared autonomous vehicles. *Transp. Res. C* 82:229–47
129. Spieser K, Samaranyake S, Frazzoli E. 2016. Vehicle routing for shared-mobility systems with time-varying demand. In *2016 American Control Conference (ACC)*, pp. 796–802. Piscataway, NJ: IEEE
130. Miller J, How JP. 2017. Predictive positioning and quality of service ridesharing for campus mobility on demand systems. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1402–8. Piscataway, NJ: IEEE
131. Ma TY, Rasulkhani S, Chow JY, Klein S. 2019. A dynamic ridesharing dispatch and idle vehicle repositioning strategy with integrated transit transfers. *Transp. Res. E* 128:417–42
132. Fielbaum A, Bai X, Alonso-Mora J. 2021. On-demand ridesharing with optimized pick-up and drop-off walking locations. *Transp. Res. C* 126:103061
133. Horni A, Nagel K, Axhausen KW. 2016. *The Multi-Agent Transport Simulation MATSim*. London: Ubiquity
134. Adnan M, Pereira FC, Azevedo CML, Basak K, Lovric M, et al. 2016. *SimMobility: a multi-scale integrated agent-based simulation platform*. Paper presented at the Transportation Research Board 95th Annual Meeting, Washington, DC, Jan. 10–14
135. Krajzewicz D, Hertkorn G, Rössel C, Wagner P. 2002. SUMO (Simulation of Urban MObility): an open-source traffic simulation. In *Proceedings of the 4th Middle East Symposium on Simulation and Modelling (MESM2002)*, ed. A Al-Akaidi, pp. 183–87. Ghent, Belg.: SCS
136. Keimer A, Bayen A. 2020. Routing on traffic networks incorporating past memory up to real-time information on the network state. *Annu. Rev. Control Robot. Auton. Syst.* 3:151–72
137. Azevedo CL, Marczuk K, Raveau S, Soh H, Adnan M, et al. 2016. Microsimulation of demand and supply of autonomous mobility on demand. *Transp. Res. Rec.* 2564:21–30

138. Auld J, Hope M, Ley H, Sokolov V, Xu B, Zhang K. 2016. Polaris: agent-based modeling framework development and implementation for integrated travel demand and network and operations simulations. *Transp. Res. C* 64:101–16
139. Jäger B, Brickwedde C, Lienkamp M. 2018. Multi-agent simulation of a demand-responsive transit system operated by autonomous vehicles. *Transp. Res. Rec.* 2672:764–74
140. Payne HJ. 1979. FREFLO: a macroscopic simulation model of freeway traffic. *Transp. Res. Rec.* 722:68–77
141. Ruch C, Hörl S, Frazzoli E. 2018. AMoDeus, a simulation-based testbed for autonomous mobility-on-demand systems. In *2018 21st IEEE International Conference on Intelligent Transportation Systems (ITSC)*, pp. 3639–44. Piscataway, NJ: IEEE
142. Hörl S. 2017. Agent-based simulation of autonomous taxi services with dynamic demand responses. *Procedia Comput. Sci.* 109:899–904
143. Miller JJJ. 2017. *Demand estimation and fleet management for autonomous mobility on demand systems*. PhD Thesis, Mass. Inst. Technol., Cambridge
144. Paull L, Tani J, Ahn H, Alonso-Mora J, Carlone L, et al. 2017. Duckietown: an open, inexpensive and flexible platform for autonomy education and research. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1497–504. Piscataway, NJ: IEEE
145. Haklay M, Weber P. 2008. OpenStreetMap: user-generated street maps. *IEEE Pervasive Comput.* 7:12–18
146. McHugh B. 2013. Pioneering open data standards: the GTFS story. In *Beyond Transparency: Open Data and the Future of Civic Innovation*, ed. B Goldstein, L Dyson, pp. 125–36. San Francisco: Code Am.
147. Donovan B, Work D. 2016. *New York City taxi trip data (2010–2013)*. Illinois Data Bank, Univ. Ill. Urbana-Champaign, deposited May 19. <https://doi.org/10.13012/J8PN93H8>
148. Matyas MB. 2020. *Investigating individual preferences for new mobility services: the case of “mobility as a service” products*. PhD Thesis, Univ. Coll. London, London
149. Kamel J, Vosooghi R, Puchinger J, Ksontini F, Sirin G. 2019. Exploring the impact of user preferences on shared autonomous vehicle modal split: a multi-agent simulation approach. *Transp. Res. Procedia* 37:115–22
150. Rossi F. 2018. *On the interaction between autonomous mobility-on-demand systems and the built environment: models and large scale coordination algorithms*. PhD Thesis, Stanford Univ., Stanford, CA
151. Gawron JH, Keoleian GA, De Kleine RD, Wallington TJ, Kim HC. 2018. Life cycle assessment of connected and automated vehicles: sensing and computing subsystem and vehicle level effects. *Environ. Sci. Technol.* 52:3249–56
152. Bonanni TM, Zardini GJ, Seccamonte F. 2021. *System and method for updating map data*. US Patent Appl. 17/129,420
153. Kim SW, Liu W, Ang MH, Frazzoli E, Rus D. 2015. The impact of cooperative perception on decision making and planning of autonomous vehicles. *IEEE Intell. Transp. Syst. Mag.* 7(3):39–50
154. Beojone CV, Geroliminis N. 2021. On the inefficiency of ride-sourcing services towards urban congestion. *Transp. Res. C* 124:102890
155. Zardini G, Censi A, Frazzoli E. 2021. Co-design of autonomous systems: from hardware selection to control synthesis. In *2021 European Control Conference (ECC)*, pp. 682–89. Piscataway, NJ: IEEE
156. Zardini G, Lanzetti N, Salazar M, Censi A, Frazzoli E, Pavone M. 2020. On the co-design of AV-enabled mobility systems. In *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. Piscataway, NJ: IEEE. <https://doi.org/10.1109/ITSC45102.2020.9294499>
157. Censi A. 2017. Uncertainty in monotone codesign problems. *IEEE Robot. Autom. Lett.* 2:1556–63
158. Ostrovsky M, Schwarz M. 2019. Carpooling and the economics of self-driving cars. In *EC ’19: Proceedings of the 2019 ACM Conference on Economics and Computation*, pp. 581–82. New York: ACM
159. Paccagnan D, Gairing M. 2021. In congestion games, taxes achieve optimal approximation. [arXiv:2105.07480 \[cs.GT\]](https://arxiv.org/abs/2105.07480)
160. Dandl F, Bogenberger K, Mahmassani HS. 2019. Autonomous mobility-on-demand real-time gaming framework. In *2019 6th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*. Piscataway, NJ: IEEE. <https://doi.org/10.1109/MTITS.2019.8883286>

161. Lanzetti N, Schiffer M, Ostrovsky M, Pavone M. 2021. On the interplay between self-driving cars and public transportation. arXiv:2109.01627 [eess.SY]
162. Psaraftis HN, Kontovas CA. 2010. Balancing the economic and environmental performance of maritime transportation. *Transp. Res. D* 15:458–62
163. Yan A, Howe B. 2020. Fairness in practice: a survey on equity in urban mobility. *Bull. IEEE Comput. Soc. Tech. Comm. Data Eng.* 42:49–63
164. Jalota D, Solovey K, Zoepf S, Pavone M. 2021. Balancing fairness and efficiency in traffic routing via interpolated traffic assignment. arXiv:2104.00098 [eess.SY]
165. Angelelli E, Morandi V, Speranza MG. 2020. Minimizing the total travel time with limited unfairness in traffic networks. *Comput. Oper. Res.* 123:105016
166. He BY, Chow JY. 2019. Optimal privacy control for transport network data sharing. *Transp. Res. Procedia* 38:792–811
167. Isaac M. 2017. How Uber deceives the authorities worldwide. *New York Times*, Mar. 3. <https://www.nytimes.com/2017/03/03/technology/uber-greyball-program-evade-authorities.html>
168. Li Y, Ouyang K, Li N, Rahmani R, Yang H, Pei Y. 2020. A blockchain-assisted intelligent transportation system promoting data services with privacy protection. *Sensors* 20:2483
169. Tsao M, Yang K, Zoepf S, Pavone M. 2021. Trust but verify: cryptographic data privacy for mobility management. arXiv:2104.07768 [cs.CR]
170. Iglesias R, Rossi F, Zhang R, Pavone M. 2019. A BCMP network approach to modeling and controlling autonomous mobility-on-demand systems. *Int. J. Robot. Res.* 38:357–74
171. Lazar DA, Byk E, Sadigh D, Pedarsani R. 2019. Learning how to dynamically route autonomous vehicles on shared roads. arXiv:1909.03664 [math.OC]
172. Gaechter J, Zanardi A, Ruch C, Frazzoli E, Pavone M. 2021. Image representation of a city and its taxi fleet for end-to-end learning of rebalancing policies. In *2021 International Conference on Robotics and Automation (ICRA)*, pp. 8076–82. Piscataway, NJ: IEEE
173. Bischoff J, Maciejewski M. 2016. Simulation of city-wide replacement of private cars with autonomous taxis in Berlin. *Procedia Comput. Sci.* 83:237–44
174. Maciejewski M, Bischoff J, Nagel K. 2016. An assignment-based approach to efficient real-time city-scale taxi dispatching. *IEEE Intell. Syst.* 31:68–77
175. Chen X, Miao F, Pappas GJ, Preciado V. 2017. Hierarchical data-driven vehicle dispatch and ride-sharing. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pp. 4458–63. Piscataway, NJ: IEEE
176. Li M, Zheng N, Wu X, Huo X. 2019. An efficient matching method for dispatching autonomous vehicles. In *2019 22nd IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 3013–18. Piscataway, NJ: IEEE
177. Braverman A, Dai JG, Liu X, Ying L. 2019. Empty-car routing in ridesharing systems. *Oper. Res.* 67:1437–52
178. Wen J, Zhao J, Jaillet P. 2017. Rebalancing shared mobility-on-demand systems: a reinforcement learning approach. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 220–25. Piscataway, NJ: IEEE
179. Spieser K, Samaranayake S, Gruel W, Frazzoli E. 2016. *Shared-vehicle mobility-on-demand systems: a fleet operator's guide to rebalancing empty vehicles*. Paper presented at the Transportation Research Board 95th Annual Meeting, Washington, DC, Jan. 10–14
180. Marczuk KA, Soh HS, Azevedo CM, Lee DH, Frazzoli E. 2016. Simulation framework for rebalancing of autonomous mobility on demand systems *MATEC Web Conf.* 81:01005