Transportation Research Part C 169 (2024) 104854

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

A dynamic macroscopic framework for pricing of ride-hailing services with an optional bus lane access for pool vehicles

Lynn Fayed, Gustav Nilsson, Nikolas Geroliminis *

Urban Transport Systems Laboratory (LUTS) École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

ARTICLE INFO

Keywords:
Macroscopic fundamental diagrams
Model predictive Control
Multimodal networks
Ride-hailing
Space allocation
Pricing

ABSTRACT

On-demand trip sharing is an efficient solution to mitigate the negative externalities e-hailing has on traffic in a network. It motivates platform operators to reduce their fleet size and serves the same demand level with a lower effective distance traveled. Users nevertheless prefer to travel solo and for shorter distances, despite the price discount they receive. By offering them the choice to pool and travel in high-occupancy dedicated bus lanes, we provide them with a larger incentive to share their rides, yet this creates additional bus delays. In this work, we develop dynamic feedback-based control policies that regulate pool vehicle access to bus lanes by adjusting the price gap between solo and pool trips, with the aim of improving multi-modal delays and providing better utilization of network capacity. First, we develop a modal- and occupancy-dependent aggregate model for private vehicles, ride-pooling, and buses based on network production, and we use this model to test different control strategies. To minimize the error between the target and actual speeds in the bus network, we design a PI controller and show that by adjusting pool trip fares, we can, with little input data, minimize this error. We also put forward a Model Predictive Control (MPC) framework to minimize the total Passenger Hours Traveled (PHT) and Waiting Times (WT) for the different travelers. Moreover, we show how the MPC framework can be utilized to impose a minimum speed in dedicated bus lanes and ensure that buses operate on schedule. The results demonstrate the possibility of improving the overall network conditions by incentivizing or discouraging pooling in the vehicle or bus network.

1. Introduction

The surge of on-demand mobility offers network commuters innovative transport alternatives for their trips. Characterized by their flexibility, convenience, and accessibility, on-demand modes have soon become widely popular, rooting their success in the fast-growing wireless communication technologies and the increasing interest in a more personalized mobility service. Ride-hailing, among many other similar modes, is nowadays a well-established transport alternative where a unique platform connects riders and drivers. Users request a ride, the request being most of the time instantaneous with no in-advance booking, and they are assigned to a nearby driver shortly after. The drivers decide whether to carry users from their origins to their destinations based on their relative evaluation of the trip attractiveness. While being as convenient as private vehicles due to their door-to-door services, these on-demand modes usually prevail over public transportation because of their shorter waiting times before pick-up (Shaheen et al., 2017).

* Corresponding author.
E-mail addresses: lynn.fayed@epfl.ch (L. Fayed), gustav.nilsson@epfl.ch (G. Nilsson), nikolas.geroliminis@epfl.ch (N. Geroliminis).

https://doi.org/10.1016/j.trc.2024.104854
Received 25 February 2024; Received in revised form 23 July 2024; Accepted 13 September 2024
Available online 25 September 2024

The burgeoning number of ride-hailing services and their wide success required authorities to impose incentive- or enforcement-based regulation strategies. These strategies are introduced to contain the negative impact that ride-hailing has on multi-modal urban traffic and user mode choice distribution. A closer look at the operation of ride-hailing services holds drivers accountable for the increase in traffic externalities (Erhardt et al., 2019). More specifically, a high number of ride-hailing drivers not only causes an increase in congestion (Diao et al., 2021) but also has a counterproductive effect by extending the waiting for users due to longer dispatching time despite the high availability of empty vehicles (Beojone and Geroliminis, 2021). Moreover, the vast majority of current ride-hailing users reported that they would use public transportation if such services were not available, hence creating a direct competition with buses (de Souza Silva et al., 2018). Clearly, the decline in the use of mass transit due to the shift in users' choice towards ride-hailing services raises multiple concerns about the undesired competition between these two modes.

To design well-informed and well-targeted policies, it is crucial to provide regulators with quantification studies that concretize the different impacts that ride-hailing has on the traffic externalities, the welfare of the drivers and riders, and the modal split between the different modes in a network (Tirachini and Gómez-Lobo, 2019). This approach guides authorities to enforce appropriate actions to prevent further propagation of these services without proper regulations. A high fleet size, for instance, shortens the waiting time of users yet increases traffic congestion in urban spaces. One way to mitigate this is through sending empty vehicles with no assigned trip to available off-street parking locations (Beojone and Geroliminis, 2021; Li et al., 2020; Xu et al., 2017) or to enforce a cap on the fleet size or the maximum allowable total Vehicle Kilometers Traveled (VKT) carried out by the ride-hailing fleet (Yu et al., 2020). Moreover, particularly when operating in a monopoly setting, ride-hailing platforms set a profit-maximizing fare without consideration of the rider's or driver's welfare. Many studies additionally investigated what the driver's wage and the rider's fare should be under a social welfare maximization framework. They argue that despite them not being sustainable from a revenue-maximization point of view, these pricing schemes are socially optimal when assessed in an equilibrium setting (Zha et al., 2016; Ke et al., 2020).

Promoting trip sharing is another strategy to reduce the VKT by drivers to serve the same demand level (Ke et al., 2021a). Trip sharing also allows the shrinking of the fleet size required to provide the same service level (Jiao and Ramezani, 2022). Generally referred to as ride-splitting or ride-pooling, this service prompts users to share their rides with other travelers in exchange for a fare discount to compensate for the extra detour incurred. Whether the passenger-to-passenger pool matching is successful depends on a myriad of factors, including the engagement levels in pooling, i.e., the willingness of users to share their rides with other users in the system, but also on their subsequent pick-up and drop-off locations. In the scope of this work, however, we ignore the possibility of a failed passenger-to-passenger pool matching, and we assume that all requests opting for pooling will eventually be pooled. We also limit the trip sharing to two passengers, noting that high-capacity on-demand micro-transit services are gaining fast momentum (Shaheen and Cohen, 2019).

Having highlighted the importance of trip sharing to alleviate the negative externalities of ride-hailing services (Tirachini, 2020), these platforms are still attracting low to moderate demand levels. Therefore, we advance in this work an occupancy- and modal-dependent allocation strategy where pooling is motivated by allowing shared trips to use dedicated bus lanes. Although the majority of impact assessment and policy evaluation studies were formulated in a static equilibrium setting, we adopt in our framework a macroscopic dynamic approach with time-varying demand to capture the non-equilibrium and transient states of the network dynamics and how the system evolves under different demand profiles during the day. More specifically, we tackle in this work the effect on multi-modal delays of incentivizing trip sharing by entitling pool users to privileged bus lane access or a higher fare discount. In the next part, we provide a detailed overview of the relevant research tackling different aspects of our work, including (i) the static and dynamic modeling of ride-hailing/ride-splitting and their service optimization, and (ii) the status of ride-hailing with respect to the other more traditional operating modes in the network.

2. Literature review

The common ground in any study tackling ride-hailing is to have a representative model capable of capturing the main features and characteristics of these services and the different stakeholders involved. The first insight into drawing the distinction between traditional taxi services and on-demand ride-hailing is to underline the appearance of a dispatching vehicle category that is non-existent in traditional taxi markets. It is the result of an online vehicle-passenger matching where the drivers' locations and the requests' origins are known to the platform (Cramer and Krueger, 2016). Nevertheless, many studies have pointed out that, despite it sometimes being useful, online location access is a source of market inefficiency in the event of a demand surge where available vehicles are quickly depleted (Castillo et al., 2017; Xu et al., 2020). A solution to this inefficiency is setting a surge pricing scheme that guarantees that the occurrence of these scenarios is avoided (Cachon et al., 2017), even if this solution raises concerns about passengers' welfare. Service pricing is therefore an important element in ride-hailing modeling, thus justifying the large body of research investigating optimal full rider fare for ride-hailing (Zha et al., 2016) and optimal discounted fare for ride-splitting (Ke et al., 2020; Zhang and Nie, 2021). Similarly, Nourinejad and Ramezani (2020) assessed service pricing but in a dynamic non-equilibrium setting with consideration of background traffic. Due to the spatial heterogeneity of demand and supply, it was also necessary to extend this framework to include an occupancy-dependent pricing scheme balancing demand and supply in multi-regions, therefore guaranteeing a more efficient service level (Zha et al., 2018; Bimpikis et al., 2019). Dynamic idle vehicle rebalancing strategies are also able to achieve the same outcomes but require having an accurate prediction of the demand in every region (Ramezani and Valadkhani, 2023; Yang and Ramezani, 2023; Guo et al., 2022).

The research lines we previously described particularly focused on modeling ride-hailing services, and very few accounted for on-demand trip sharing in their framework. This is mainly because microscopically modeling trip sharing, which is one of the main

contributions of this work, is complex to conduct, especially when the number of passengers participating in a trip exceeds two. The focus hence deviated towards finding some empirical and universal laws for driver and passenger detours (Tachet et al., 2016) or assessing the different factors influencing the quality of a pooled trip (Soza-Parra et al., 2023). Moreover, the passenger-to-passenger matching and the vehicle dispatching require advanced exact algorithms or heuristics to solve them in real-time settings (Jung et al., 2016; Alisoltani et al., 2021; Alonso-Mora et al., 2017), even if in some work on two passenger-pooling, passenger-to-passenger matching probability prediction returned similar results to simulation settings (Wang et al., 2021). Consequently, this computational effort makes it complex to integrate the matching with upper-level optimization problems like vehicle rebalancing or dynamic lane usage, as in our case.

Positioning ride-hailing attractiveness relative to public transportation leads to questioning whether these two services are complimentary or substitutionary. In areas where public transit is well-developed, ride-hailing is rather viewed as a first/last-mile solution complementing bus or metro services (Hall et al., 2018). However, this does not eliminate the potential competition between the two modes (Mo et al., 2021), where, in many cases, ride-hailing substitutes transit, hence causing an inevitable increase in the total VKT (Tirachini and Gómez-Lobo, 2019). This observation led many researchers to formulate user equilibrium under different available alternatives where users have the choice to use ride-hailing either for their full trips or for a subpart of their trips (Ke et al., 2021b; Zhu et al., 2021; Ma et al., 2023). They showed that subsidizing ride-hailing as a first/last-mile solution can indeed reinforce modal complementarity, despite reducing ride-hailing profit.

The purpose of this work is to assess the importance of ride-hailing in a multi-modal context and to advance an adaptive pricing strategy that, combined with an adequate vehicle allocation scheme, allows us to minimize total delays in the network. This occupancy- or modal-dependent vehicle allocation framework has been previously studied in the context of High-Occupancy Toll (HOT) lanes (Toledo et al., 2017; Cohen et al., 2022) or dedicated lanes for Autonomous Vehicles (AV) on the link level (Lamotte et al., 2017), or for buses on the network level (Tsitsokas et al., 2021). To the best of our knowledge, however, no work on modal- and occupancy-dependent allocation strategies to minimize overall network delays has considered ride-hailing services in its framework. In our previous work (Fayed et al., 2023b), this strategy has been assessed under fixed demand and static settings, where we focused on studying equilibrium properties under the proposed allocation strategy, and we showed that the optimal modal allocation is demand-dependent. Nevertheless, since dynamically adjusting network space is practically infeasible, we put forward in this work a dynamic model that captures the demand fluctuations through price control policies in fixed vehicle allocation settings.

The contribution of this work is twofold. First, we build a macroscopic multi-modal dynamic model that utilizes Macroscopic Fundamental Diagrams (MFD) to determine the dynamics of private vehicles, bus users, and ride-hailing services (Geroliminis and Daganzo, 2008). We assume that ride-hailing users can choose to travel solo in the vehicle network or to pool in the bus network. In contrast to Fayed et al. (2023a), we additionally consider the option of pooling in the vehicle network. Second, we introduce a regulatory pricing scheme for the pooling options and use the dynamic model developed within a control framework to determine what are the additional discount or fare that must be allocated to pool users to steer the system towards its optimum. In other words, we propose a solution that minimizes the Total Time Spent (TTS) by all mode users in the network through adjusting the fares for pooled trips. Moreover, to guarantee minimal disturbances to buses by pool ride-hailing trips in bus lanes, we impose a minimum speed in the bus network. This step ensures that buses continue to operate on schedule such that the utilization of their network by pool ride-hailing vehicles does not cause them significant delays. We highlight that the purpose of this study is not to mitigate traffic externalities at the expense of buses, but to make use of the underutilized bus network capacity when the overall demand conditions allow. Ultimately, the aim is to focus on the potential improvement of available capacity utilization by ride-hailing vehicles in multi-modal networks. The remainder of the paper is organized as follows. In the following section, Section 3, we describe the modal- and occupancy-dependent space allocation strategy we present in this work and elaborate on the macroscopic state dynamics for the different transportation modes using the network under assessment. Next, in Section 4, we lay out the different control frameworks used for the purpose of improving network delays. To demonstrate the performance of the different controllers, we display the results of the simulations we run in Section 5. The paper finally concludes with the main findings and future directions in Section 6.

3. Macroscopic multi-modal framework

In the following section, we describe our macroscopic framework by delineating the modal-dependent allocation strategy we propose. Next, we put forward the aggregate network traffic production model according to the vehicle allocation policy advanced and use it to define the traffic dynamics for the different transportation modes under consideration.

3.1. Modal-dependent vehicle allocation

In this work, we assess a modal-dependent allocation strategy where users travel in different parts of the network based on the travel mode they choose to adopt.

Assumption 1. In the network under consideration, travelers perform their trips using one of the available transportation alternatives in the set $\mathcal{M}$: private vehicles pv, buses b, and ride-hailing services rs such that $\mathcal{M} := \{pv, b, rs\}$.

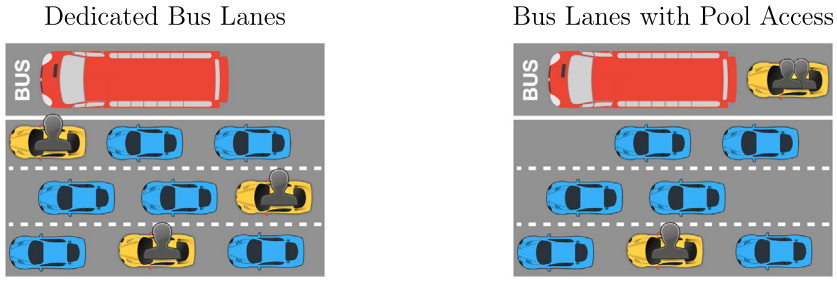

Fig. 1. A sketch reflecting the network configuration without our allocation strategy (left), and the configuration with pool access to bus lanes (right)..

The dynamics for each alternative are modeled in discrete time with time step $k \in \mathcal{K} := \{0, \dots, k_{\max}\}$, and the duration of each time step is $\tau > 0$. Moreover, we let $\bar{\mathcal{K}} := \mathcal{K} \setminus \{k_{\max}\}$. Every transportation mode under consideration has an exogenous and time-dependent demand $Q_j(k)$ for $j \in \mathcal{M}$ expressed in passengers per hour. Although a comprehensive demand approach for multi-modal user choice under the proposed framework is crucial to capture the inter-modal choice, we limit our analysis to intra-modal interaction in this work. A more general logit model capturing other mode choices would require additional real data to be estimated, and it can be a research priority in the future. Commuters who opt for the ride-hailing alternative choose to either travel solo or pool their trips with other users of the service. In terms of capacity utilization, private vehicles perform their trips in the subspace of the network occupying a fixed fraction $\alpha \in [0, 1]$ of the total network space. Buses from their sides solely travel in dedicated bus lanes spanning over a fraction $1 - \alpha$ of the total space. We denote the vehicle and bus networks by $\mathcal{V}$ and $\mathcal{B}$ respectively, such that the set of available subnetworks is $\mathcal{N} := \{\mathcal{V}, \mathcal{B}\}$. Note that we roughly assume a uniform mixture of network space such that the vast majority of the road network has a spatial fraction allocated for buses. The vehicle trip routing is therefore nearly the same, irrespective of whether it is performed in the vehicle or bus network. Finally, if ride-hailing users opt for traveling solo, then the entirety of their trip is performed in the vehicle network $\mathcal{V}$. If, however, ride-hailing users choose to pool, they are granted the opportunity to either travel in the vehicle network $\mathcal{V}$ or the bus network $\mathcal{B}$.

Assumption 2. The ride-hailing driver fleet $N > 0$ is constant and time-independent in our framework, and the different drivers can exclusively be in one of the following states:

(i) driving an empty vehicle in the vehicle network $\mathcal{V}$, where the total number of drivers in this state is denoted by n_e,
(ii) delivering a solo trip in the vehicle network $\mathcal{V}$, where the total number of drivers in this state is denoted by n_s,
(iii) delivering a pool trip in the vehicle network $\mathcal{V}$, where the total number of drivers in this state is denoted by $n_p^{\mathcal{V}}$ and,
(iv) delivering a pool trip in the bus network $\mathcal{B}$, where the total number of drivers in this state is denoted by $n_p^{\mathcal{B}}$.

It should be noted that we consider all of the aggregate states to be non-negative and continuous.

To sum up, the proposed modal-dependent allocation strategy splits the overall network space into a vehicle network $\mathcal{V}$ and a bus network $\mathcal{B}$, each occupied by different categories of vehicles. The vehicle network $\mathcal{V}$ is utilized by private vehicles, empty, solo, and a portion of pooled ride-hailing vehicles. Parallely, the bus network $\mathcal{B}$ is used by buses and the remaining portion of pooled ride-hailing vehicles. Moreover, irrespective of the type of ride-hailing trips, vehicle dispatching always occurs in the vehicle network $\mathcal{V}$, and pool vehicles in $\mathcal{B}$ are only allowed to physically switch to the bus network after picking up the first passenger in the trip. A visualization of this occupancy-dependent allocation policy is shown in Fig. 1.

Moreover, let n_{pv} be the number of private vehicles in the vehicle network $\mathcal{V}$ and n_b the number of buses in the bus network $\mathcal{B}$. Having defined the different vehicle categories, we know that the ride-hailing fleet at any time step $k \in \mathcal{K}$, under the assumption of a fixed fleet size, has to satisfy $N = n_e(k) + n_s(k) + n_p^{\mathcal{V}}(k) + n_p^{\mathcal{B}}(k)$. Note that in some urban spaces, parking spots are available for idle vehicles that are not assigned a trip. However, in this work, we assume that idle vehicles always cruise in the network when unassigned. The accumulation in the vehicle network $n_{\mathcal{V}}$ at any time step $k \in \mathcal{K}$ is $n_{\mathcal{V}}(k) = n_{pv}(k) + n_e(k) + n_s(k) + n_p^{\mathcal{V}}(k)$, and the accumulation in the bus network $n_{\mathcal{B}}$ at any time step $k \in \mathcal{K}$ is $n_{\mathcal{B}}(k) = n_b + n_p^{\mathcal{B}}(k)$. The number of operating buses in $\mathcal{B}$ is time-independent in our framework. A summary of all the different variables used in this work, along with their descriptions, is displayed in Table 1.

3.2. Aggregate traffic flow model

In the following part, we elaborate on the aggregate traffic model we use to estimate the speed in the vehicle and bus networks, i.e., $\mathcal{V}$ and $\mathcal{B}$. Let $P : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ denote the full network production function without any dedicated bus lanes. A production function is a metric of network performance. It measures the total distance traveled by the different vehicles in the network per unit of time. The total network production P is dependent on the full network accumulation n, and can be calculated using the running network speed $v : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $P(n) = nv(n)$. Following Ni and Cassidy (2019), Sirmatel et al. (2021), we can compute the production $P_{\mathcal{V}} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ in the vehicle network and the production $P_{\mathcal{B}} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ in the bus network using the space

Table 1
List of notations.

Variable name	Description
$\mathcal{M}$	Set of transportation modes where $\mathcal{M} = \{pv, b, rs\}$
$\mathcal{N}$	Set of available subnetworks
Q_j	Demand for mode j, $j \in \mathcal{M}$
n	Total vehicle accumulation
P	Total production in the network
v	Network speed
α	Fraction of space allocated to the vehicle network $\mathcal{V}$
$n_{\mathcal{V}}$, $n_{\mathcal{B}}$	Accumulation in the vehicle and bus network $\mathcal{V}$ and $\mathcal{B}$ respectively
n_{pv}, n_b	Number of private vehicles and buses respectively
n_e, n_s	Number of empty and solo ride-hailing vehicles respectively
$n_p^{\mathcal{V}}$, $n_p^{\mathcal{B}}$	Number of pool ride-hailing vehicles in network $\mathcal{V}$ and $\mathcal{B}$ respectively
$P_{\mathcal{V}}$, $P_{\mathcal{B}}$	Production in the vehicle and bus network $\mathcal{V}$ and $\mathcal{B}$ respectively
P_p, P_b	Production of pool vehicles and buses respectively in the bus network $\mathcal{B}$
$v_{\mathcal{V}}$, $v_{\mathcal{B}}$	Speed in the vehicle and bus network $\mathcal{V}$ and $\mathcal{B}$ respectively
v_p, v_b	Speed of pool vehicles and buses respectively in the bus network $\mathcal{B}$
$\bar{t}_d$, $\bar{s}$	Average bus dwell time and spacing between stops respectively
o_b	Average bus occupancy
k	Macro-simulation time step with duration τ
N	Ride-hailing fleet size
O_{pv}, O_b	Trip completion rate for private vehicles and bus passengers respectively
O_s, $O_p^{\mathcal{V}}$, $O_p^{\mathcal{B}}$	Trip completion rate for solo, pool vehicles in $\mathcal{V}$, and pool vehicles in $\mathcal{B}$, respectively
M	Matching rate between empty ride-hailing vehicles and waiting requests
a_0, α_e, α_c	Cobb–Douglas matching function parameters
c	Number of waiting ride-hailing requests to be matched
A	Number of abandoning ride-hailing requests
$w_{\max}$	Average maximum waiting tolerance of requests before abandoning
U_s, $U_p^{\mathcal{V}}$, $U_p^{\mathcal{V}}$	Utilities for solo, pooling in $\mathcal{V}$, and pooling in $\mathcal{B}$ alternatives, respectively
u_s, $u_p^{\mathcal{V}}$, $u_p^{\mathcal{V}}$	Utilities for solo, pooling in $\mathcal{V}$, and pooling in $\mathcal{B}$ alternatives, respectively, excluding the decision variables
β_s, $\beta_{\mathcal{V}}$, $\beta_{\mathcal{B}}$	Fraction of the ride-hailing demand opting for solo, pooling in $\mathcal{V}$ and pooling in $\mathcal{B}$, respectively
μ	Mode choice scale parameter
κ	Value of time
Δl_p, Δl_d	Ride-hailing passenger and driver detour respectively
$\bar{l}_{pv}$, $\bar{l}_s$, $\bar{l}_b$	Average trip length for private vehicles, solo ride-hailing, and bus trips, respectively
$\phi_{\mathcal{V}}$, $\phi_{\mathcal{B}}$	Decision variables reflecting the fare increase/decrease for pooling in $\mathcal{V}$ and $\mathcal{B}$ respectively
$\xi_{\mathcal{V}}$, $\xi_{\mathcal{B}}$	Variables function of $\phi_{\mathcal{V}}$ and $\phi_{\mathcal{B}}$ respectively
F_s, F_p	Ride-hailing fare for traveling solo or pooling respectively
$\tilde{F}_s$, $\tilde{F}_p^{\mathcal{V}}$, $\tilde{F}_p^{\mathcal{B}}$	Fare for traveling solo, pooling in $\mathcal{V}$, or pooling in $\mathcal{B}$, respectively, including the decision variables
K_p, K_i	Proportional and integral gain for the PI controller respectively
$\bar{v}_b$	Desired speed of buses in $\mathcal{B}$
ϵ	Error term between the desired and actual speed of buses

allocation factor α such that $P_{\mathcal{V}}(\alpha n) = \alpha P(n)$ and $P_{\mathcal{B}}(\bar{\alpha} n) = \bar{\alpha} P(n)$ where $\bar{\alpha} = 1 - \alpha$. This is because the production of the different subnetworks consists of a rescaling of the production of the original network according to the factor α. Note that for simplicity, we drop in this Section 3.2 the time-dependencies of the production, speed, and accumulation variables.

The relationship between production and accumulation is valid as long as the vehicles commuting in the bus network are standard vehicles. However, since buses perform frequent stops to board and alight passengers, they slow down or even halt the pool ride-hailing vehicles utilizing the same space. To capture this interaction between buses and pooling vehicles utilizing the bus network, we partition the production in the bus network into pool vehicle production $P_p : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and bus production $P_b : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, each dependent on both the number of pooling vehicles $n_p^{\mathcal{B}}$ and the number of buses n_b to obtain a 3D-MFD. More information on how the 3D-MFD is empirically observed and validated in the literature can be found in Geroliminis et al. (2014) and Loder et al. (2017).

Analogous to the reasoning behind the split of the production function, the same applies here to the aggregate estimation of the vehicular speed in the vehicle network $v_{\mathcal{V}} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and vehicular speed in the bus network $v_{\mathcal{B}} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ where $v_{\mathcal{V}}(\alpha n) = v(n)$ and $v_{\mathcal{B}}(\bar{\alpha} n) = v(n)$. Having estimated the individual speed function for every subnetwork, the production value in the vehicle network $\mathcal{V}$ is hence given by $P_{\mathcal{V}}(n_{\mathcal{V}}) = n_{\mathcal{V}} v_{\mathcal{V}}(n_{\mathcal{V}})$ and that in the bus network $\mathcal{B}$ is given by $P_{\mathcal{B}}(n_{\mathcal{B}}) = n_{\mathcal{B}} v_{\mathcal{B}}(n_{\mathcal{B}})$. However, for the latter function, we account for the influence of buses on the vehicle speed by reducing $v_{\mathcal{B}}$ with a factor $r(n_b)$ that is dependent on the number of buses in the network where $r : \mathbb{R}_{\geq 0} \to (0, 1]$ and $\frac{\mathrm{d}r}{\mathrm{d}n_b} < 0$. Therefore, the actual speed of pool vehicles, that we denote by $v_p : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, is given by $v_p(n_p^{\mathcal{B}}, n_b) = v_{\mathcal{B}}(n_{\mathcal{B}}) r(n_b)$. Regarding the bus operating speed that we denote $v_b : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, it must account for the repetitive stops that buses perform at stations such that

$$v_b(n_p^{\mathcal{B}}, n_b) = \left(\frac{1}{1 + v_p(n_p^{\mathcal{B}}, n_b) \frac{\bar{t}_d}{\bar{s}}} \right) v_p(n_p^{\mathcal{B}}, n_b), \tag{1}$$

where $\bar{t}_d$ and $\bar{s}$ are the average bus dwell time and the spacing between bus stops respectively. The objective of the factor $\frac{1}{1+v_p\left(n_p^{\mathcal{B}},n_b\right)\frac{\bar{t}_d}{\bar{s}}} \leq 1$ is to convert the bus running speed to the bus operating speed. This is because theoretically, the bus speed v_b in urban spaces potentially matches v_p if no stopping at stations is considered. To elaborate more, the value $1 + v_p\left(n_p^{\mathcal{B}}, n_b\right)\frac{\bar{t}_d}{\bar{s}}$ returns the additional time factor required by buses to cross the same distance as regular vehicles, including stopping at stations. Therefore, when the running speed $v_p\left(n_p^{\mathcal{B}}, n_b\right)$ is multiplied by the reciprocal of this factor, the result is the adjusted operating speed expressed in unitary distance. It follows that the pool vehicle production in $\mathcal{B}$ is $P_p\left(n_p^{\mathcal{B}}, n_b\right) = n_p^{\mathcal{B}} v_p\left(n_p^{\mathcal{B}}, n_b\right)$ and the bus production is $P_b\left(n_p^{\mathcal{B}}, n_b\right) = n_b v_b\left(n_p^{\mathcal{B}}, n_b\right)$. Note that pool ride-hailing vehicles also perform repetitive stops to pick up and/or drop off passengers, but the duration and frequency of these stops are negligible compared to buses, and their stops have minimal interference with traffic flow. To summarize, the speed $v_{\mathcal{B}}$ returns the theoretical speed in the network $\mathcal{B}$ if all vehicles consist of regular ride-hailing vehicles. Given that buses are the main utilizers of this network, this speed is then reduced by a factor of $r < 1$ to capture the influence of buses on pool ride-hailing traffic. This returns the running speed in the bus network v_p. This speed is then further reduced to obtain the operating bus speed v_b which also includes the dwell time of buses at stops.

The aggregate network-dependent production functions are used to estimate the trip completion rate – or outflow – for every category of vehicles under consideration. While this approximation does not provide the same level of accuracy compared to trip-based models, it still allows for a tractable analysis compared to the latter which is too complex for this sort of application (Sirmatel et al., 2021). As a result, in the following subsections, we utilize the production-based multi-modal macroscopic traffic model to estimate the changes as a function of time of private vehicle accumulation, ride-hailing fleet assignment, and bus occupancy.

3.3. Network dynamics

Previously, we have defined the modal-dependent vehicle allocation policy and the subnetwork-dependent macroscopic traffic functions. In the following part, we elaborate on the aggregate dynamics of private vehicles, ride-hailing fleet, and bus average occupancies according to the proposed allocation policy. Starting with the private vehicle category, the change in accumulation n_{pv} follows from the flow conservation and is given by

$$n_{pv}(k+1) = n_{pv}(k) + \tau\left[\frac{Q_{pv}(k)}{\bar{o}_{pv}} - O_{pv}(k)\right], \forall k \in \bar{\mathcal{K}}, \tag{2}$$

where $\bar{o}_{pv} > 0$ is the average occupancy of a private vehicle and $O_{pv}(k)$ is the trip completion rate computed using

$$O_{pv}(k) = \frac{n_{pv}(k)}{n_{\mathcal{V}}(k)}\frac{P_{\mathcal{V}}\left(n_{\mathcal{V}}(k)\right)}{\bar{l}_{pv}}, \tag{3}$$

where $\bar{l}_{pv} > 0$ is the constant average trip distance between the origin–destination pairs of private vehicle users. Note that we assume a homogeneous mixture of private and ride-hailing vehicles. The first part of (3) hence considers the fraction of the total vehicle network $\mathcal{V}$ belonging to private vehicles, while the second part computes the trip completion rate in $\mathcal{V}$ of private vehicles.

Moving to the ride-hailing mode, the various available options have different travel costs. Based on the relative cost of each option, the users choose to travel solo in $\mathcal{V}$, to pool in $\mathcal{V}$, or to pool in $\mathcal{B}$. Therefore, we let U_s denote the disutility for traveling solo, $U_p^{\mathcal{V}}$ the disutility for pooling in the vehicle network, and $U_p^{\mathcal{B}}$ the disutility for pooling in bus lanes. Under the assumption of user homogeneity, the expressions of the different utilities at time step $k \in \mathcal{K}$ are

$$U_s(k) = \tilde{F}_s(k) + \kappa\frac{\bar{l}_s}{v_{\mathcal{V}}\left(n_{\mathcal{V}}(k)\right)}, \tag{4a}$$

$$U_p^{\mathcal{V}}(k) = \tilde{F}_p^{\mathcal{V}}(k) + \kappa\frac{\bar{l}_s + \Delta l_p}{v_{\mathcal{V}}\left(n_{\mathcal{V}}(k)\right)}, \tag{4b}$$

$$U_p^{\mathcal{B}}(k) = \tilde{F}_p^{\mathcal{B}}(k) + \kappa\frac{\bar{l}_s + \Delta l_p}{v_p\left(n_p^{\mathcal{B}}(k), n_b\right)}, \tag{4c}$$

where $\kappa > 0$ is the value of time for homogeneous ride-hailing users. The variable $\bar{l}_s > 0$ is the average trip length for a solo trip, and $\Delta l_p \geq 0$ is the pool detour distance that passengers incur in case they opt for pooling. The detour is the same irrespective of the network chosen due to the uniform split of network space assumption in Section 3.1. The variable $\tilde{F}_s$ is the fare for traveling solo in $\mathcal{V}$, $\tilde{F}_p^{\mathcal{V}}$ is the fare for pooling in $\mathcal{V}$, and $\tilde{F}_p^{\mathcal{B}}$ is the fare for pooling in $\mathcal{B}$. In this work, we assume that $\tilde{F}_s(k)$ is constant such that $\tilde{F}_s(k) = F_s$ for all $k \in \mathcal{K}$ where $F_s > 0$ is the solo trip fare set by the platform operator. On the contrary, if $F_p > 0$ is the pool trip fare set by the operator, then $\tilde{F}_p^{\mathcal{V}}(k) = F_p + \phi_{\mathcal{V}}(k)$ and $\tilde{F}_p^{\mathcal{B}}(k) = F_p + \phi_{\mathcal{B}}(k)$ where $\phi_{\mathcal{V}}(k) \in \mathbb{R}$ and $\phi_{\mathcal{B}}(k) \in \mathbb{R}$ are the control fares for pooling in the vehicle network $\mathcal{V}$ and bus network $\mathcal{B}$ respectively. The control fares are introduced to steer the total network towards different objectives that we expand on in Section 4. We note here that despite the second pool ride-hailing passenger incurring a larger waiting time before pick-up, this inconvenience is implicitly captured within the detour. It is therefore the reason why the waiting time is dropped from the different cost functions.

The relative values of each of the disutility functions are used to compute the modal share for every available ride-hailing alternative. Therefore, let $\beta_{\mathcal{V}} \in [0, 1]$ and $\beta_{\mathcal{B}} \in [0, 1]$ be the fraction of the total ride-hailing demand that will choose to pool

in $\mathcal{V}$ and $\mathcal{B}$ respectively at time step $k \in \mathcal{K}$, then to model ride-hailing user's choice, we resort to the traditional and widely-used multinomial logit model as in Zhang and Nie (2021), Vignon et al. (2023), and we obtain that

$$\beta_{\mathcal{V}}(k) = \frac{e^{-\mu U_p^{\mathcal{V}}(k)}}{e^{-\mu U_s(k)} + e^{-\mu U_p^{\mathcal{V}}(k)} + e^{-\mu U_p^{\mathcal{B}}(k)}}, \tag{5a}$$

$$\beta_{\mathcal{B}}(k) = \frac{e^{-\mu U_p^{\mathcal{B}}(k)}}{e^{-\mu U_s(k)} + e^{-\mu U_p^{\mathcal{V}}(k)} + e^{-\mu U_p^{\mathcal{B}}(k)}}, \tag{5b}$$

where $\mu > 0$ is the scale parameter. The portion of users choosing to go solo at time step $k \in \mathcal{K}$ is $\beta_s(k)$ such that $\beta_s(k) = 1 - \beta_{\mathcal{V}}(k) - \beta_{\mathcal{B}}(k)$.

For ease of implementation purposes, we reformulate the utility and mode choice functions to set apart the control and state variables. Consequently, if we redefine the disutilities of the three available ride-hailing alternatives, excluding the controllable price changes, by u_s, $u_p^{\mathcal{V}}$, and $u_p^{\mathcal{B}}$, we get that

$$u_s(k) = F_s + \kappa \frac{\bar{l}_s}{v_{\mathcal{V}}\left(n_{\mathcal{V}}(k)\right)}, \tag{6a}$$

$$u_p^{\mathcal{V}}(k) = F_p + \kappa \frac{\bar{l}_s + \Delta l_p}{v_{\mathcal{V}}\left(n_{\mathcal{V}}(k)\right)}, \tag{6b}$$

$$u_p^{\mathcal{B}}(k) = F_p + \kappa \frac{\bar{l}_s + \Delta l_p}{v_p\left(n_p^{\mathcal{B}}(k), n_b\right)}, \tag{6c}$$

where $u_p^{\mathcal{V}}$ for instance is different from $U_p^{\mathcal{V}}$ because it does not consider the control fare $\phi_{\mathcal{V}}$ in its formulation. If we set $\xi_{\mathcal{V}}(k)$ and $\xi_{\mathcal{B}}(k)$ as two variables that are function of $\phi_{\mathcal{V}}(k)$ and $\phi_{\mathcal{B}}(k)$ for all $k \in \mathcal{K}$, we get

$$\xi_{\mathcal{V}}(k) := e^{-\mu\phi_{\mathcal{V}}(k)}, \tag{7}$$

$$\xi_{\mathcal{B}}(k) := e^{-\mu\phi_{\mathcal{B}}(k)}. \tag{8}$$

Consequently, we can rewrite $\beta_{\mathcal{V}}(k)$ and $\beta_{\mathcal{B}}(k)$ as follows

$$\beta_{\mathcal{V}}(k) = \frac{\xi_{\mathcal{V}}(k)e^{-\mu u_p^{\mathcal{V}}(k)}}{e^{-\mu u_s(k)} + \xi_{\mathcal{V}}(k)e^{-\mu u_p^{\mathcal{V}}(k)} + \xi_{\mathcal{B}}(k)e^{-\mu u_p^{\mathcal{B}}(k)}}, \tag{9}$$

$$\beta_{\mathcal{B}}(k) = \frac{\xi_{\mathcal{B}}(k)e^{-\mu u_p^{\mathcal{B}}(k)}}{e^{-\mu u_s(k)} + \xi_{\mathcal{V}}(k)e^{-\mu u_p^{\mathcal{V}}(k)} + \xi_{\mathcal{B}}(k)e^{-\mu u_p^{\mathcal{B}}(k)}}. \tag{10}$$

We conclude that if $c(k)$ is the number of customers waiting to be matched at time k for all $k \in \mathcal{K}$, then we know that the number of passengers choosing to travel solo is $c_s(k) = \left(1 - \beta_{\mathcal{V}}(k) - \beta_{\mathcal{B}}(k)\right)c(k)$ and the number of passengers choosing to pool is $\left(\beta_{\mathcal{V}}(k) + \beta_{\mathcal{B}}(k)\right)c(k)$. By resorting to a Cobb–Douglas meeting function as in Zha et al. (2016), Nourinejad and Ramezani (2020), we compute the matching rate $M(k)$ at time step $k \in \mathcal{K}$ using

$$M(k) = a_0 n_e(k)^{\alpha_e} \left(c_s(k) + \frac{1}{2}c_p(k)\right)^{\alpha_c}, \tag{11}$$

where $a_0 > 0$, $\alpha_e > 0$, and $\alpha_c > 0$ are the Cobb–Douglas meeting function parameters. Note that a factor $\frac{1}{2}$ is added to model that every single pool trip consists of two passengers. This formulation does not imply that pool passengers will share the complete trip together but rather that every two passengers opting for pooling are assigned to a single empty ride-hailing vehicle. Moreover, we assume full driver compliance with the matching decision, such that drivers will not reject any assignment by the platform. Subsequently, we can compute the dynamics of empty vehicles n_e at any time step using

$$\begin{aligned} n_e(k+1) = n_e(k) + \tau \Bigg[& \frac{n_s(k)}{n_{\mathcal{V}}(k)} \frac{P_{\mathcal{V}}\left(n_{\mathcal{V}}(k)\right)}{\bar{l}_s} + \\ & \frac{n_p^{\mathcal{V}}(k)}{n_{\mathcal{V}}(k)} \frac{P_{\mathcal{V}}\left(n_{\mathcal{V}}(k)\right)}{\bar{l}_s + \Delta l_d} + \frac{P_p\left(n_p^{\mathcal{B}}(k), n_b\right)}{\bar{l}_s + \Delta l_d} - M(k) \Bigg], \quad \forall k \in \bar{\mathcal{K}}. \end{aligned} \tag{12}$$

where the first three elements of (12) represent the completion rate of solo trips $O_s(k) = \frac{n_s(k)}{n_{\mathcal{V}}(k)} \frac{P_{\mathcal{V}}\left(n_{\mathcal{V}}(k)\right)}{\bar{l}_s}$, the completion rate of pool trips in $\mathcal{V}$, $O_p^{\mathcal{V}}(k) = \frac{n_p^{\mathcal{V}}(k)}{n_{\mathcal{V}}(k)} \frac{P_{\mathcal{V}}\left(n_{\mathcal{V}}(k)\right)}{\bar{l}_s + \Delta l_d}$, and the completion rate of pool trips in $\mathcal{B}$, $O_p^{\mathcal{B}}(k) = \frac{P_p\left(n_p^{\mathcal{B}}(k), n_b\right)}{\bar{l}_s + \Delta l_d}$. The last element of (12) denotes the number of empty vehicles that are matched and have therefore exited this category. The variable $\Delta l_d \geq 0$ represents the driver detour, which is the additional distance traveled by drivers to perform a pool trip relative to a solo one. The driver detour is different from the passenger detour Δl_p, which is defined as the additional distance traveled by passengers to pick up and/or drop off another passenger.

Moving to the discretized dynamics of the solo vehicle category n_s, we, by following the conservation of flow, get that

$$n_s(k+1) = n_s(k) + \tau \left[\beta_s(k)M(k) - \frac{n_s(k)}{n_{\mathcal{V}}(k)} \frac{P_{\mathcal{V}}\left(n_{\mathcal{V}}(k)\right)}{\bar{l}_s}\right], \quad \forall k \in \bar{\mathcal{K}}. \tag{13}$$

Similarly, the discretized dynamics for the pool vehicle in $\mathcal{V}$ category, $n_p^{\mathcal{V}}$, are given by

$$n_p^{\mathcal{V}}(k+1) = n_p^{\mathcal{V}}(k) + \tau\left[\beta_{\mathcal{V}}(k)M(k) - \frac{n_p^{\mathcal{V}}(k)}{n_{\mathcal{V}}(k)}\frac{P_{\mathcal{V}}\left(n_{\mathcal{V}}(k)\right)}{\bar{l}_s + \Delta l_d}\right], \quad \forall k \in \bar{\mathcal{K}}. \tag{14}$$

Finally, the discretized dynamics for the pool vehicles in $\mathcal{B}$, category, $n_p^{\mathcal{B}}$, are given by

$$n_p^{\mathcal{B}}(k+1) = n_p^{\mathcal{B}}(k) + \tau\left[\beta_{\mathcal{B}}(k)M(k) - \frac{P_p\left(n_p^{\mathcal{B}}(k), n_b\right)}{\bar{l}_s + \Delta l_d}\right], \quad \forall k \in \bar{\mathcal{K}}. \tag{15}$$

Because the total ride-hailing fleet size is constant, the sum of the inflow to (13), (14), and (15) is equal to the outflow of (12), and vice versa. Note that in this analysis, we assume that pool passengers opting for a trip in the vehicle $\mathcal{V}$ or bus network $\mathcal{B}$ are assigned to separate idle ride-hailing drivers.

To further illustrate how the vehicle allocation policy operates, we consider the following example. Once a pool trip in $\mathcal{B}$ is assigned to an empty ride-hailing vehicle driving in $\mathcal{V}$, the latter drives to pick up the first passenger in the vehicle network $\mathcal{V}$, then switches to the bus network $\mathcal{B}$ to perform the rest of the trip before moving back to the vehicle network once it delivers the last passenger in the vehicle.

So far, we have defined the changes in the ride-hailing vehicles category. In a similar manner, the changes in the number of passengers in the queue waiting to be assigned are

$$c(k+1) = c(k) + \tau\left[Q_{rs}(k) - \left(1 + \beta_{\mathcal{V}}(k) + \beta_{\mathcal{B}}(k)\right)M(k)\right] - A(k), \quad \forall k \in \bar{\mathcal{K}}. \tag{16}$$

The element $\left(1 + \beta_{\mathcal{V}}(k) + \beta_{\mathcal{B}}(k)\right)M(k)$ in (16) represents the outflow of the waiting requests category. It accounts for the fact that a pool trip match requires subtracting two passengers from this category compared to a solo trip match. Therefore, if all ride-hailing users opt for pooling, we have that $\left(1 + \beta_{\mathcal{V}}(k) + \beta_{\mathcal{B}}(k)\right) = 2$, and twice the matching rate is subtracted from the queue of waiting requests to account for two-passenger trip sharing. Else, if no ride-hailing users opt for pooling, we have that $\left(1 + \beta_{\mathcal{V}}(k) + \beta_{\mathcal{B}}(k)\right) = 1$, and the matching rate is subtracted from the queue of waiting request to account for single-passenger trips. Finally, the variable $A(k) \geq 0$ represents the number of cancellation/abandoning requests due to extended waiting periods before pick-up caused by limited empty ride-hailing vehicle availability. Therefore, if the average maximum waiting tolerance of ride-hailing customers is given by $w_{\max}$, then we can estimate the number of abandoning requests using

$$A(k) = \max\left(c(k) - \frac{1}{k}\sum_{\tilde{k}=1}^{k} M(\tilde{k})w_{\max}, 0\right), \quad \forall k \in \mathcal{K}. \tag{17}$$

The abandonment approximation in (17) first computes, at any time step k, what should be the queue length if the requests' maximum waiting tolerance value is equal to $w_{\max}$. This is done by multiplying the matching rate $M(k)$ by the exogenous parameter $w_{\max}$. However, given that the matching rate changes with time, we approximate its value by taking the average over all the previous time steps. This way, we can get an estimation of the maximum expected queue length irrespective of when the ride-hailing request appeared in the network. When this quantity is then subtracted from the actual queue length at time step k, $c(k)$, we get an approximation of the number of requests that left the system due to long waiting times. Note that since this value is only an estimate of the abandonment, the max operator is added to guarantee that the abandonment does not exceed the actual queue length of waiting requests.

The dynamics of the third transportation mode, buses, are formulated in terms of changes in occupancy rather than changes in the number of vehicles. This is because we assume that the number of buses n_b traveling in the bus network $\mathcal{B}$ is constant and their occupancy is time-dependent. Although some variations in the number of buses are expected over the course of the day, we consider an analysis period during which this number is roughly constant. Assuming a uniform bus occupancy o_b over the available fleet of buses n_b, the discretized dynamics of o_b are given by

$$o_b(k+1) = o_b(k) + \frac{\tau}{n_b}\left[Q_b(k) + \frac{A(k)}{\tau} - \frac{P_b\left(n_p^{\mathcal{B}}(k), n_b\right)}{\bar{l}_b}o_b(k)\right], \quad \forall k \in \bar{\mathcal{K}}, \tag{18}$$

where $\bar{l}_b$ is the average trip length by bus. Note here that the abandoning ride-hailing requests at time step $k \in \mathcal{K}$, $A(k)$, are considered here as an additional demand for the bus occupancy category. The last term of (18) represents the bus passenger trip completion rate $O_b(k) = \frac{P_b\left(n_p^{\mathcal{B}}(k), n_b\right)}{\bar{l}_b}o_b(k)$.

A summary of the proposed multi-modal vehicle allocation modeling framework is provided in Fig. 2.

In the next section, Section 4, we elaborate on the ride-hailing pricing scheme that we set forth to reduce the multi-modal user delays by utilizing the dynamic model we previously described.

4. Control framework

The objective of the vehicle allocation model we advance in this work is to redistribute the network space over the available transportation modes. However, the choice of ride-hailing users to travel solo or to pool is not always aligned with the objective of improving the TTS in the network or ensuring that buses will stay on schedule. Therefore, in this section, we develop a regulatory

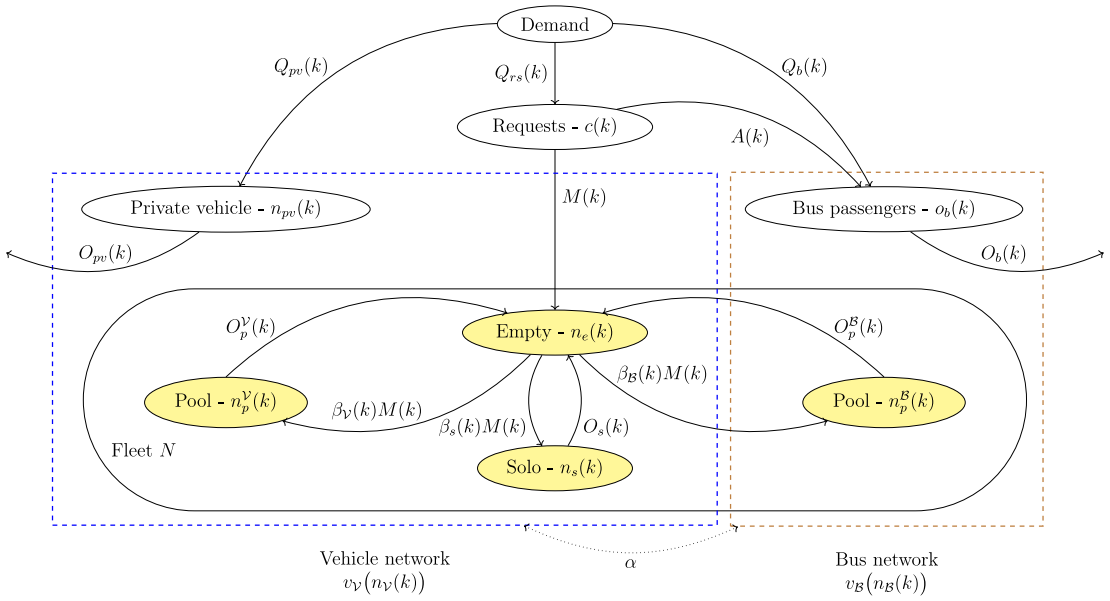

Fig. 2. A schematic sketch of the model under consideration. Private vehicles, empty ride-hailing vehicles, solo trips, and a portion of pool ride-hailing trips drive in the vehicle subnetwork $\mathcal{V}$, whereas buses and the remaining pool trip portion perform their trips in the bus subnetwork $\mathcal{B}$. The nodes in yellow correspond to the different categories of ride-hailing vehicles. The blue dashed rectangle encloses all the vehicles utilizing the vehicle subnetwork $\mathcal{V}$ while the brown dashed rectangle encloses all the vehicles utilizing the bus subnetwork $\mathcal{B}$..

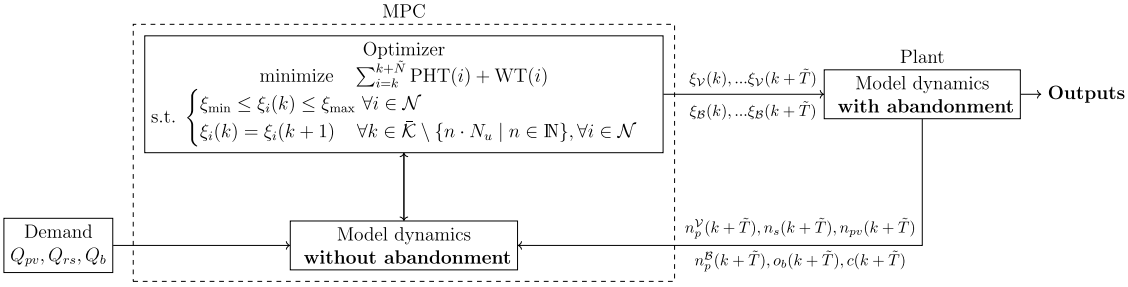

Fig. 3. Implementation of the MPC framework with abandonment. In this framework, the MPC dynamics run without abandonment with a prediction horizon of $\tilde{N}$ whereas the plant dynamics run with abandonment. After every update step $\tilde{T}$, we reinitialize the MPC with some updated initial states retrieved from the plant dynamics.

pricing scheme where pool ride-hailing users receive a discount or pay an extra fee for the purpose of ensuring that this fraction of pool trips in bus lanes does not worsen multi-modal user delays. To do so, we evaluate two different control policies. The first one is a Proportional–Integral (PI) controller with the objective of keeping the bus network at a certain speed to not delay the bus operation. The PI controller is myopic and only requires information about the speed in the bus network to compute the control prices. Its implementation is consequently straightforward because it does not need prior knowledge of the future demand but rather only requires the definition of a bus speed set point. The second control policy we evaluate is the Model Predictive Control (MPC) with the objective of minimizing the total travel time.

4.1. PI control

As opposed to more elaborate controllers, the ease of implementation of the PI allows for a baseline comparison between different pricing control schemes. A PI framework provides hence a sufficient step forward to guarantee that the proposed vehicle allocation policy is not a hurdle to bus circulation. Because our proposed policy moves pool vehicles to bus lanes, it is crucial to ensure that the disturbances to bus users are minimized, all while improving the travel time for the remaining travelers. We do so by changing the values of the pool trip control fares $\phi_{\mathcal{B}}$ according to the difference between the actual bus speed in network $\mathcal{B}$ and the target bus speed, which we denote by $\bar{v}_b > 0$. This difference is also referred to as the bus speed error. The choice of $\phi_{\mathcal{B}}$ as our control variable in the PI framework is justified by the direct effect this quantity has on the amount of trip pooling in the bus network. In addition, we keep track of the previous errors for the last $N_e \in \mathbb{N}$ time steps, where N_e represents the number of preceding time

steps over which the error is computed. As a consequence, if $\epsilon(k) := \left(\bar{v}_b - v_b(n_p^{\mathcal{B}}(k), n_b)\right)$ defines the error term at time step k, then the expression for $\phi_{\mathcal{B}}$ at time step $k \in \mathcal{K}$ is

$$\phi_{\mathcal{B}}(k) = K_p \epsilon(k) + \frac{K_i}{N_e} \sum_{\tilde{k}=\max(k-(N_e+1),0)}^{k-1} \epsilon(\tilde{k}), \tag{19}$$

where $K_p > 0$ and $K_i \geq 0$ are the constant proportional gain and integral gain. We point out here that, in particular under the PI formulation, we directly control $\phi_{\mathcal{B}}$ rather than $\xi_{\mathcal{B}}$, because having the price fare as the control input is straightforward. In the result section, we show how the two different parameters, K_p and K_i, affect the performance of the PI controller by bridging the gap between the actual and target bus speeds. The drawback of the PI controller, however, is that it only takes into account the speed of buses irrespective of the other mode users. Next, we show how the MPC implementation is capable of circumventing this matter.

4.2. Model predictive control

Unlike the PI controller, the aim of the MPC framework is to improve the network TTS or delays using our proposed policy. We quantify delays using the total Passenger Hours Traveled (PHT) of multi-modal users at any time step $k \in \mathcal{K}$, here equal to $\text{PHT}(k) = \tau\left[n_{pv}(k)\bar{o}_{pv} + n_b o_b(k) + n_s(k) + \left(n_p^{\mathcal{V}}(k) + n_p^{\mathcal{B}}(k)\right)\bar{o}_p\right]$, and the total Waiting Time (WT) of ride-hailing requests at time step $k \in \mathcal{K}$, which is given by $\text{WT}(k) = \tau c(k)$. Therefore, the formulation of the MPC framework is given by

$$\begin{aligned}
\underset{\xi_i \in \mathbb{R}^{\tilde{N}},\, i \in \mathcal{N}}{\text{minimize}} \quad & \sum_{j=k}^{k+\tilde{N}} \text{PHT}(j) + \text{WT}(j) \\
\text{subject to} \quad & \xi_i(j) \in [\xi_{\min}, \xi_{\max}] && \forall j \in [k, k+\tilde{N}],\ \forall i \in \mathcal{N} \\
& \xi_i(j) = \xi_i(j+1) && \forall j \in [k, k+\tilde{N}-1] \setminus \{n \cdot N_u \mid n \in \mathbb{N}\},\ \forall i \in \mathcal{N} \\
& A(j) = 0 && \forall j \in [k, k+\tilde{N}] \\
& (2), (12), (13), (14), (15), (18)
\end{aligned} \tag{20}$$

where $\xi_{\min}$ and $\xi_{\max}$ are the exogenous lower and upper bounds of the control variable, $\bar{o}_p$ is the average occupancy of a pool trip such that $\bar{o}_p \in (1, 2]$, and $\tilde{N}$ is the prediction horizon. The second constraint of (20) makes sure that the control actions are only updated every $N_u \in \mathbb{N}$ time steps. This avoids frequent fluctuations in the service pricing of ride-hailing, which is desirable from a user perspective. The third constraint sets the number of abandoning ride-hailing requests to 0. This is because the abandoning function in (17) increases the computational complexity of the MPC framework to an extent where the problem cannot be handled by the solver used. The computational complexity is increased because the abandonment function (17) is not only dependent on the previous state but all earlier states as well, which makes it harder to solve the optimization problem. Moreover, abandonment constitutes the source of uncertainty in our framework, and it is not straightforward to estimate the exact shape of the function. For these reasons, we run the MPC dynamics without abandonment, and we reinitialize the state variables with the actual state dynamics with abandonment produced by the plant, i.e., n_{pv}, n_s, $n_p^{\mathcal{V}}$, $n_p^{\mathcal{B}}$, c, and o_b after every $\tilde{T}$ time steps, $\tilde{T}$ being the update horizon in our framework. A more detailed explanation of this approach is provided in Fig. 3.

5. Numerical study

In the following section, we perform a numerical study to evaluate the impact of our proposed allocation scheme, and we show the influence of different control strategies on the overall network performance. To do so, we describe in Section 5.1 the simulation parameters that we utilize for illustration purposes. Next, in Section 5.2, we assess network delays for scenarios without and with abandonment. In Sections 5.3 and 5.4, we delve more into the results for the different pool price controllers implemented in this work.

5.1. Macro-simulation parameters

Next, we describe the simulation environment that we implement to test the potential benefits of our proposed strategy along with the different control schemes. For this purpose, we consider a network characterized by an MFD that aggregates the microscopic traffic dynamics. Its functional form is given by the production function $P(n) = A_0 n^3 + B_0 n^2 + C_0 n$, $A_0 = 5.74 \cdot 10^{-9}$, $B_0 = -1.02 \cdot 10^{-3}$, and $C_0 = 36$ for $n \in [0, 58536]$. The production function is calibrated to fit the MFD in (Beojone and Geroliminis, 2021). The fractional split α partitioning the full network space into a vehicle network and a bus network has a value equal to 0.8 and is usually an unalterable property of the infrastructure. The value of α yields two production functions for every subnetwork according to the relationship described in Section 3. Therefore, having the full network production function $P(n)$, we can obtain the subnetwork production functions by setting $P_{\mathcal{V}}(n_{\mathcal{V}}) = \alpha P(n)$, and $P_{\mathcal{B}}(n_{\mathcal{B}}) = \bar{\alpha} P(n)$. Similarly, given the full network speed $v(n) = A_0 n^2 + B_0 n + C_0$ for $n \in [0, 58536]$, we can obtain the subnetwork speed functions $v_{\mathcal{V}}(n_{\mathcal{V}}) = v(n)$ and $v_{\mathcal{B}}(n_{\mathcal{B}}) = v(n)$, where $n_{\mathcal{V}} = \alpha n$ and $n_{\mathcal{B}} = \bar{\alpha} n$. Transforming the obtained vehicle speed MFD, $v_{\mathcal{B}}$, into a three-dimensional bus MFD requires the multiplication of the speed with a reduction factor $r(n_b)$ such that $r(n_b) = e^{-6.5 \cdot 10^{-4} n_b}$. This function provides a reasonable decrease to the speed in $\mathcal{B}$ due to the existence

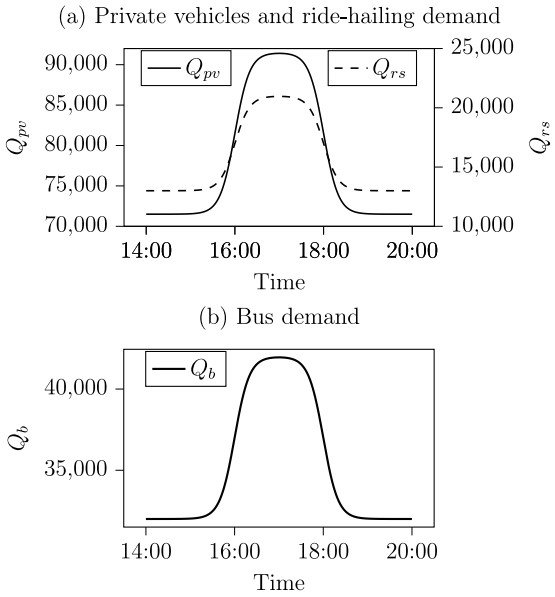

Fig. 4. Time-dependent multi-modal demand profile in pax/h.

of buses. The resulting speed function described in Section 3.2 allows the computation of the pool vehicles running speed in the bus network v_p, such that $v_p(n_B, n_b) = v_B(n_B)r(n_b)$. Similarly, the bus operating speed is straightforwardly obtained by factoring in the spacing $\bar{s}$ between bus stations and the dwell time $\bar{t}_d$ at stops where $\bar{s} = 0.8$ km, and $\bar{t}_d = 30$ s in (1). With respect to the Cobb–Douglas meeting function that we adopt in (11), its constant parameters are equal to 0.025, 0.93, and 0.98 for a_0, α_e, and α_c respectively, with the same order of magnitude as in Nourinejad and Ramezani (2020). To estimate the outflow for every category of vehicle, we assume that the average trip length for private vehicles is equal to that of solo trips such that $\bar{l}_{pv} = \bar{l}_s = 3.86$ km, which is the average trip length retrieved from Beojone and Geroliminis (2021). Since bus trips are generally longer than direct origin–destination trips, we consider that the average bus trip distance is $\bar{l}_b = 1.4\bar{l}_{pv}$. Similarly, pool trips are also longer than $\bar{l}_s$ due to the additional detour that passengers/drivers have to incur. The value of this detour is generally dependent on the number of passengers willing to engage in pooling. For the scope of this analysis, however, we will assume that the detour remains constant such that $\Delta l_d = 0.7\bar{l}_s$ and $\Delta l_p = 0.15\bar{l}_s$. The integration of a demand-dependent detour value will be examined in future works. The average occupancy of pool vehicles is $\bar{o}_p = 1.5$, as the driver is excluded from the occupancy count. With respect to the private vehicles' occupancy $\bar{o}_{pv}$, we use a value of 1.2. The ride-hailing request waiting tolerance $w_{\max}$ is set to 15 min.

Moving to the multinomial logit model dictating the choice between solo and pooled trips, we consider a mode choice scale parameter $\mu = 1$, and a value of time $\kappa = 30$ CHF/h. In real-world settings, the value of these parameters can be retrieved and calibrated using real data to accurately reflect user choices. The static trip fares for solo F_s and pool F_p are 5 and 4 CHF, respectively. We note here that these values usually change with the total ride-hailing demand. Since the objective of this study is to determine the values of $\phi_{\mathcal{V}}$ and $\phi_{\mathcal{B}}$ irrespective of the values of F_s and F_p, we discard the demand-dependent basic fare variations. The fleet size N remains constant over the full simulation framework and is set to 3500 vehicles. The simulation runs over six hours and is discretized such that the duration of every time step τ is equal to 6 s. The demand profile for private vehicles and ride-hailing is displayed in Fig. 4a where the increase in demand starts at around 16:00 before it goes back to its original value at around 18:00. The demand for buses varies in a similar manner, as also shown in Fig. 4b. This demand profile is synthetic, and is designed for illustrative purposes such that the provided examples showcase a realistic scenario.

5.2. Macro-simulation results for the total multi-modal delays

In the following section, we present the multi-modal delays for different network configuration settings to better assess the importance of the proposed vehicle allocation strategy, and the different pricing control schemes presented in this work. To achieve this objective, we first compare the occupancy-dependent vehicle allocation scheme with settings where the choices for the pooling users are limited under no control schemes. Likewise, we additionally compare different pool pricing control configurations. Particularly, we define the following no control and control scenarios that we assess: Scenario 1 considers a solo ride-hailing services where all trips are performed in the vehicle network $\mathcal{V}$. Scenario 2 is when ride-hailing services offer the pooling option exclusively in the vehicle network $\mathcal{V}$ while Scenario 3 is when ride-hailing services offer the pooling option exclusively in the bus network $\mathcal{B}$. Scenario 4 is an extreme scenario where all ride-hailing trips are pooling and utilizing the bus network $\mathcal{B}$. Scenario 5 replicates the model presented in Section 3 where ride-hailing users have the option to either travel solo in $\mathcal{V}$, pool in $\mathcal{V}$, or pool in $\mathcal{B}$. Moving to Scenarios 6–10, they represent configurations with control. More specifically, Scenario 6 refers to the case where $\phi_{\mathcal{B}}$ is controlled

Table 2
Macro-simulation results without abandonment.

#	Scenario		PHT+WT	PHT	WT
	Vehicle network	Bus network	[pax.h]	[pax.h]	[pax.h]
1	No pool ($\beta_{\mathcal{V}} = 0$)	No pool ($\beta_{\mathcal{B}} = 0$)	246 980	193 913	53 067
2	Optional pool ($\beta_{\mathcal{V}} \in [0, 1]$)	No pool ($\beta_{\mathcal{B}} = 0$)	239 819	198 116	41 703
3	No pool ($\beta_{\mathcal{V}} = 0$)	Optional pool ($\beta_{\mathcal{B}} \in [0, 1]$)	190 872	176 442	14 430
4	No pool ($\beta_{\mathcal{V}} = 0$)	Forced pool ($\beta_{\mathcal{B}} = 1$)	223 210	188 097	35 113
5	Optional pool ($\beta_{\mathcal{V}} \in [0, 1]$)	Optional pool ($\beta_{\mathcal{B}} \in [0, 1]$)	188 954	177 389	11 565
6	PI control – $\phi_{\mathcal{B}}$		191 383	178 709	12 674
7	MPC – $\phi_{\mathcal{B}}$		188 316	176 721	11 595
8	MPC ($\underline{v}_b$) – $\phi_{\mathcal{B}}$		191 388	178 477	12 911
9	MPC – $\phi_{\mathcal{V}}$ and $\phi_{\mathcal{B}}$		187 030	176 348	10 682
10	MPC ($\underline{v}_b$) – $\phi_{\mathcal{V}}$ and $\phi_{\mathcal{B}}$		189 662	178 314	11 348

by the PI as per Section 4.1, while Scenarios 7 and 9 replicate the MPC framework as per Section 4.2, with one control variable $\phi_{\mathcal{B}}$, and two control variables $\phi_{\mathcal{B}}$ and $\phi_{\mathcal{V}}$, respectively. Finally, to have a basis for comparison between the PI and MPC frameworks, we introduce Scenarios 8 and 10 to be the MPC implementation, with the nuance of adding a penalty to the cost function when the bus speed goes below a desired value, i.e., when we introduce a soft constraint on the bus speed. We let this lower bound on the bus speed, which we denote with $\underline{v}_b$, be equal to the PI set point $\bar{v}_b$, and penalize the violation in the objective function of the MPC that is now composed of the PHT, the WT, and the weighted degree of violation of the bus speed. We do so to guarantee that the bus services do not significantly lose performance, even if this implies a worse total PHT and WT for the overall network. This modification is integrated into the MPC settings in Scenarios 8 and 10, with the sole difference between the two being that the former has a single control variable $\phi_{\mathcal{B}}$ while the latter has two control variables $\phi_{\mathcal{B}}$ and $\phi_{\mathcal{V}}$. Note that since the abandonment function reflects a rough approximation of the waiting tolerance of ride-hailing users, we present the results for all the previously described scenarios for the case where the abandonment function $A(k) = 0$, i.e., the case without abandonment, and the case where the abandonment function $A(k)$ is given according to (17), i.e., the case with abandonment. By doing so, we will examine how robust the different control strategies are to users' behaviors when it comes to abandonment.

The different TTS for the no abandonment case are presented in Tables 2 and 3, and in Tables 4 and 5 for the case with abandonment. The major difference between the two cases is that the PHT and WT values are higher for the no abandonment case, and this is because ride-hailing users have to wait for a long time before being served. Irrespective of the case under consideration, among the no control scenarios, the worst performing scenario is when no pooling is involved, i.e., when both $\beta_{\mathcal{V}}$ and $\beta_{\mathcal{B}}$ are equal to zero (Scenario 1). This is because, for constant fleet sizes, solo travel results in longer queues and longer waiting times, especially when all ride-hailing vehicles are occupied and few vehicles are available for pick-up. For the case with abandonment, Table 4 shows that this scenario also has the highest number of abandoning requests. When pooling is only allowed in the vehicle network, i.e., $\beta_{\mathcal{B}} = 0$ and $\beta_{\mathcal{V}} \in [0, 1]$ or equivalently Scenario 2, the TTS are much greater than the scenario where pooling is optional and only allowed in the bus network with $\beta_{\mathcal{V}} = 0$ (Scenario 3). This is due to users opting for pooling in the bus network, causing ride-hailing vehicles to travel at a higher speed and making them available soon after to perform a new trip. We note here that we also provide the simulation results for when $\beta_{\mathcal{B}} = 1$ which implies $\beta_{\mathcal{V}} = 0$ (Scenario 4), meaning that all ride-hailing users are pooling in the bus network. The reason for that is to show that this extreme solution causes significant travel times for bus users and long waiting times for ride-hailing users, and is therefore not attractive at the system level. Motivated by these observations, the need to regulate the share of each ride-hailing alternative becomes more substantiated. For this purpose, we resort to different controllers to find a proper pricing scheme that minimizes the observed TTS.

Moving to Scenario 5 emulating the vehicle-dependent allocation policy presented in this work, the changes of the state variables presented in Section 3 for the network under consideration without any regulatory interventions are shown in Fig. 5. Under this scenario, the choice of users is solely dictated by the platform-set fares and the subnetworks' travel time. Note that we display here the simulation results with no abandonment, i.e., when $w_{\max}$ is infinitely large or equivalently $A(k) = 0$ for all $k \in \mathcal{K}$, and with abandonment according to the model presented in Section 3.3. Both results are presented to assess the implications of the approximated abandonment functions in (17) on the macro-simulation variables. During peak hours, the accumulation of private vehicles n_{pv} spiraled up in Fig. 5a. With respect to the different states of the ride-hailing fleet size, it can be observed from Figs. 5b–5d, that the number of solo trip vehicles, n_s, and pool trip in $\mathcal{B}$ vehicles, $n_p^{\mathcal{B}}$, increased whereas the number of pool trip in $\mathcal{V}$ vehicles, $n_p^{\mathcal{V}}$, remained almost constant during peak hours. Filtering out the effect of the abandonment function, it is clearly the case that the results with abandonment show a higher bus occupancy, and a lower number of solo vehicles n_s and pool vehicles in $\mathcal{V}$ $n_p^{\mathcal{V}}$, where the difference between the two curves is more accentuated during peak hours. The deterioration of the condition in the vehicle network, reflected in the reduction in speed $v_{\mathcal{V}}$ in Fig. 5f, prompts users to pool in the bus network. This is confirmed by looking at the increase in the share of pool users opting to travel in the bus network $\beta_{\mathcal{B}}$ and the decrease in the share of pool users opting to travel in the bus network $\beta_{\mathcal{V}}$ as seen in 5i. The TTS for all users in the network for this specific scenario, i.e., when $\beta_{\mathcal{V}}$ and $\beta_{\mathcal{B}}$ are determined through (9) and (10) respectively, are displayed in Tables 2 and 4, where the sum of PHT and WT is equal to 188954 pax.h for the case without abandonment, and 185157 pax.h with 4732 abandoning requests in the case with abandonment.

Table 3
Multi-modal TTS without abandonment.

#	Scenario		PHT_{pv}	PHT_{rs}	PHT_{b}	$n_b \int \max(\bar{v}_b - v_b, 0)dt$
	Vehicle network	Bus network	[pax.h]	[pax.h]	[pax.h]	[km]
1	No pool ($\beta_V = 0$)	No pool ($\beta_B = 0$)	114 856	18 518	60 539	0
2	Optional pool ($\beta_V \in [0,1]$)	No pool ($\beta_B = 0$)	114 856	22 721	60 539	0
3	No pool ($\beta_V = 0$)	Optional pool ($\beta_B \in [0,1]$)	84 623	20 770	71 049	2634
4	No pool ($\beta_V = 0$)	Forced pool ($\beta_B = 1$)	75 462	26 713	85 922	11 869
5	Optional pool ($\beta_V \in [0,1]$)	Optional pool ($\beta_B \in [0,1]$)	87 991	21 572	67 825	958
6	PI control – ϕ_B		89 527	21 798	67 384	90
7	MPC – ϕ_B		87 355	21 564	67 802	1430
8	MPC ($\underline{v}_b$) – ϕ_B		91 039	21 627	65 811	2
9	MPC – ϕ_V and ϕ_B		87 669	21 186	67 493	1354
10	MPC ($\underline{v}_b$) – ϕ_V and ϕ_B		91 258	21 452	65 604	2

Table 4
Macro-simulation results with abandonment.

#	Scenario		PHT+WT	Abandonment
	Vehicle network	Bus network	[pax.h]	[pax]
1	No pool ($\beta_V = 0$)	No pool ($\beta_B = 0$)	211 898	17 816
2	Optional pool ($\beta_V \in [0,1]$)	No pool ($\beta_B = 0$)	210 505	15 970
3	No pool ($\beta_V = 0$)	Optional pool ($\beta_B \in [0,1]$)	185 664	4812
4	No pool ($\beta_V = 0$)	Forced pool ($\beta_B = 1$)	196 565	12 630
5	Optional pool ($\beta_V \in [0,1]$)	Optional pool ($\beta_B \in [0,1]$)	185 157	4732
6	PI control – ϕ_B		186 471	5398
7	MPC – ϕ_B		184 691	4665
8	MPC ($\underline{v}_b$) – ϕ_B		186 559	5435
9	MPC – ϕ_V and ϕ_B		184 884	3979
10	MPC ($\underline{v}_b$) – ϕ_V and ϕ_B		186 799	4050

Table 5
Multi-modal PHT.

#	Scenario		PHT_{pv}	PHT_{rs}	PHT_{b}
	Vehicle network	Bus network	[pax.h]	[pax.h]	[pax.h]
1	No pool ($\beta_V = 0$)	No pool ($\beta_B = 0$)	114 856	17 815	65 578
2	Optional pool ($\beta_V \in [0,1]$)	No pool ($\beta_B = 0$)	114 855	21 367	65 058
3	No pool ($\beta_V = 0$)	Optional pool ($\beta_B \in [0,1]$)	84 902	19 811	72 407
4	No pool ($\beta_V = 0$)	Forced pool ($\beta_B = 1$)	77 072	24 039	87 451
5	Optional pool ($\beta_V \in [0,1]$)	Optional pool ($\beta_B \in [0,1]$)	88 282	20 504	69 169
6	PI control – ϕ_B		89 558	20 587	69 086
7	MPC – ϕ_B		88 494	20 479	68 444
8	MPC ($\underline{v}_b$) – ϕ_B		91 852	20 408	66 887
9	MPC – ϕ_V and ϕ_B		88 409	20 637	68 335
10	MPC ($\underline{v}_b$) – ϕ_V and ϕ_B		91 602	20 720	66 658

5.3. PI controller framework

Having compared the total objective function for the different scenarios, we delve next into the details of the numerical application of the PI controller. To guarantee that our vehicle allocation strategy does not worsen the situation for mainly bus users, we implement the PI control framework in our simulation and report the different state variables in Fig. 6 and 7. The choice of the set point for the desired speed in the bus network, however, remains complex because bus users should ideally travel at the highest possible speed, and this speed is defined by the bus operator or the traffic regulators. In the results provided, we set $\bar{v}_b$ to 17 km/h while the default bus speed in the absence of cars is $v_b(0, n_b) = 19$ km/h. This choice of set point ensures that the permissible decrease in bus speeds due to the bus lane usage remains within acceptable ranges.

The plots in Figs. 6 and 7 show the system dynamics for different proportional and integral gain values K_p and K_i respectively, all with $N_e = 100$ time steps, for the cases without abandonment and with abandonment. The reason for displaying the different graphs for both approaches is to check that despite the integration of an approximated abandonment function, the PI implementation is still robust and performs in an equal manner. The variations of n_{pv}, n_s, $n_p^{\mathcal{B}}$, and $n_p^{\mathcal{V}}$ in the no abandonment case in Figs. 6a–6e are almost similar to what is observed in Figs. 5a–5e, except for the lower number of solo trips and higher number of pool trips in $\mathcal{B}$, especially during off-peak period. This is because the bus network is significantly underutilized during off-peak hours, and its capacity allows

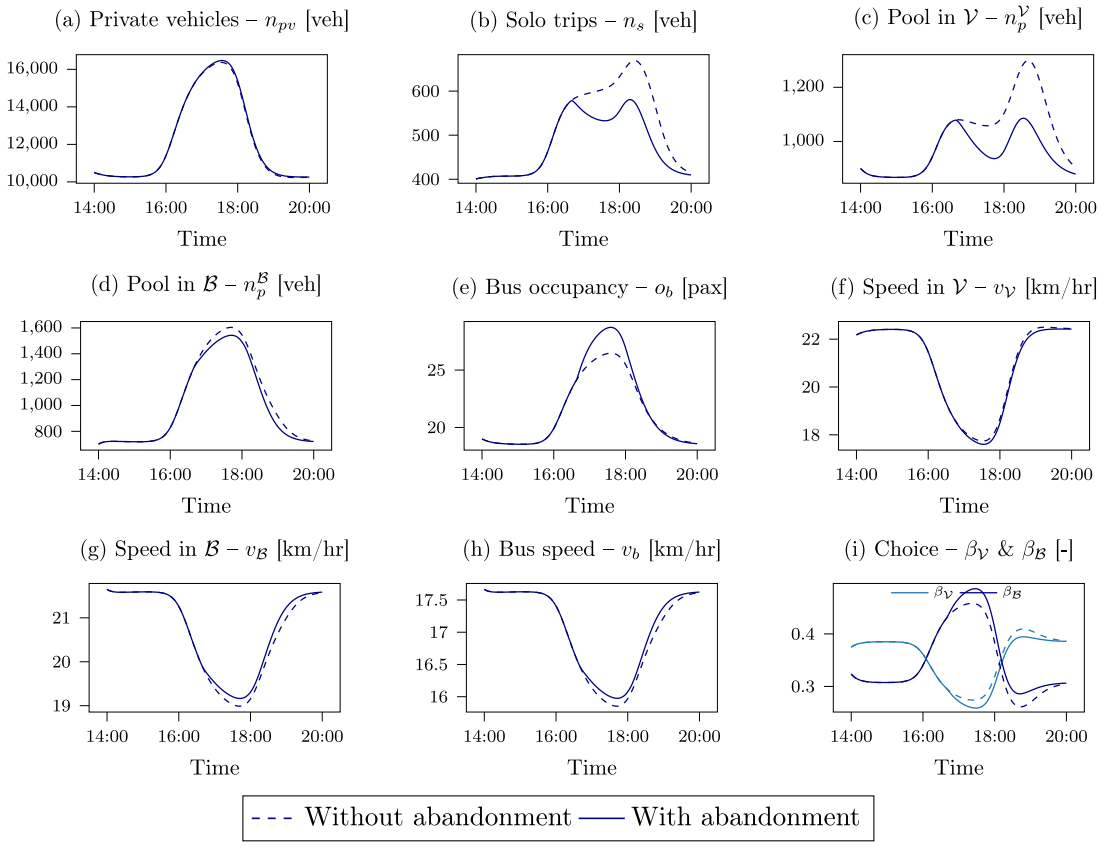

Fig. 5. Time-dependent model variables for the no control scenario 5 where $\beta_{\mathcal{V}} \in [0, 1]$ and $\beta_{\mathcal{B}} \in [0, 1]$ without and with abandonment for (a) private vehicle accumulation, (b) solo trip ride-hailing vehicles, (c) pool trip ride-hailing accumulation in $\mathcal{V}$, (d) pool trip ride-hailing accumulation in $\mathcal{B}$, (e) average bus occupancy, (f) speed in the vehicle network $\mathcal{V}$, (g) vehicle speed in the bus network $\mathcal{B}$, (h) bus speed in the bus network $\mathcal{B}$, and (i) fraction of pool trip in $\mathcal{V}$ and $\mathcal{B}$ respectively.

some ride-hailing users to pool their trips in $\mathcal{B}$. This justifies the exorbitantly high discount $\phi_{\mathcal{B}}$ granted to users at the beginning of the simulation in Fig. 6j to encourage them to pool their trips in $\mathcal{B}$. Note here that in the PI framework, we focus on controlling $\phi_{\mathcal{B}}$ because it has a direct influence on the bus speed v_b. In other words, we set $\phi_{\mathcal{V}}$ to 0 in our numerical experiment. The observations for the scenario with abandonment in Fig. 7 are similar, with Fig. 7k additionally illustrating that the 4665 abandoning requests in Table 4 mainly occur during peak hours. Since the PI implementation does not directly account for the user delays, this justifies the similarities in the variation of the price variable $\phi_{\mathcal{B}}$ in Figs. 6j and 7j for cases without and with abandonment, respectively.

Moreover, we observe that picking reasonable values for K_p and K_i yields realistic pricing scenarios, all while achieving the desired objective of bridging the gap between the actual and target bus speeds, as can be seen in Fig. 6j compared to scenarios with no integral term. Due to the time-dependent nature of the demand in our simulations, the PI controller fails to achieve lower objective function values compared to the scenario in which the vehicle allocation strategy advanced in this work is implemented with no control, i.e. (Scenario 5), where the TTS are equal to 191383 and 188954 pax.h, respectively for the case without abandonment, and 185157 and 186471 pax.h for the case with abandonment.

The approach we have adopted so far accounts for bus user delays without consideration of the overall network performance. Therefore, the previous choice of set point does not guarantee a convenient solution for all network users. In Figs. 8 and 9, we display the values of PHT and WT for various PI set points and different on-peak period bus demand multipliers for the case with and without abandonment, respectively. The displayed results reflect that a lower $\bar{v}_b$ is acceptable if the objective is to minimize multi-modal TTS. A more practical way to determine this network operational point is to extend the static model developed in (Fayed et al., 2023b), and compute the value of bus speed at optimality for the pair of time-invariant private vehicles and bus demands. The results of this approach are shown in Fig. 10. We note here that this static approach does not have any consideration of abandonment in its framework. From the figure, we can see that for the current setting, the set-point is demand-dependent. The corresponding minimum PHT values for the different demand combinations are shown in Fig. 11. Since bus lanes usually occupy a small fraction of the network infrastructure, increasing bus demand has the most significant impact on the PHT values.

Fig. 12 shows the variation of the PHT results for different choices of the PI controller set points for the bus demand profile in Fig. 4b, and compares them with the optimal set point found by solving the static formulation. Clearly, the value of $\bar{v}_b$ minimizing

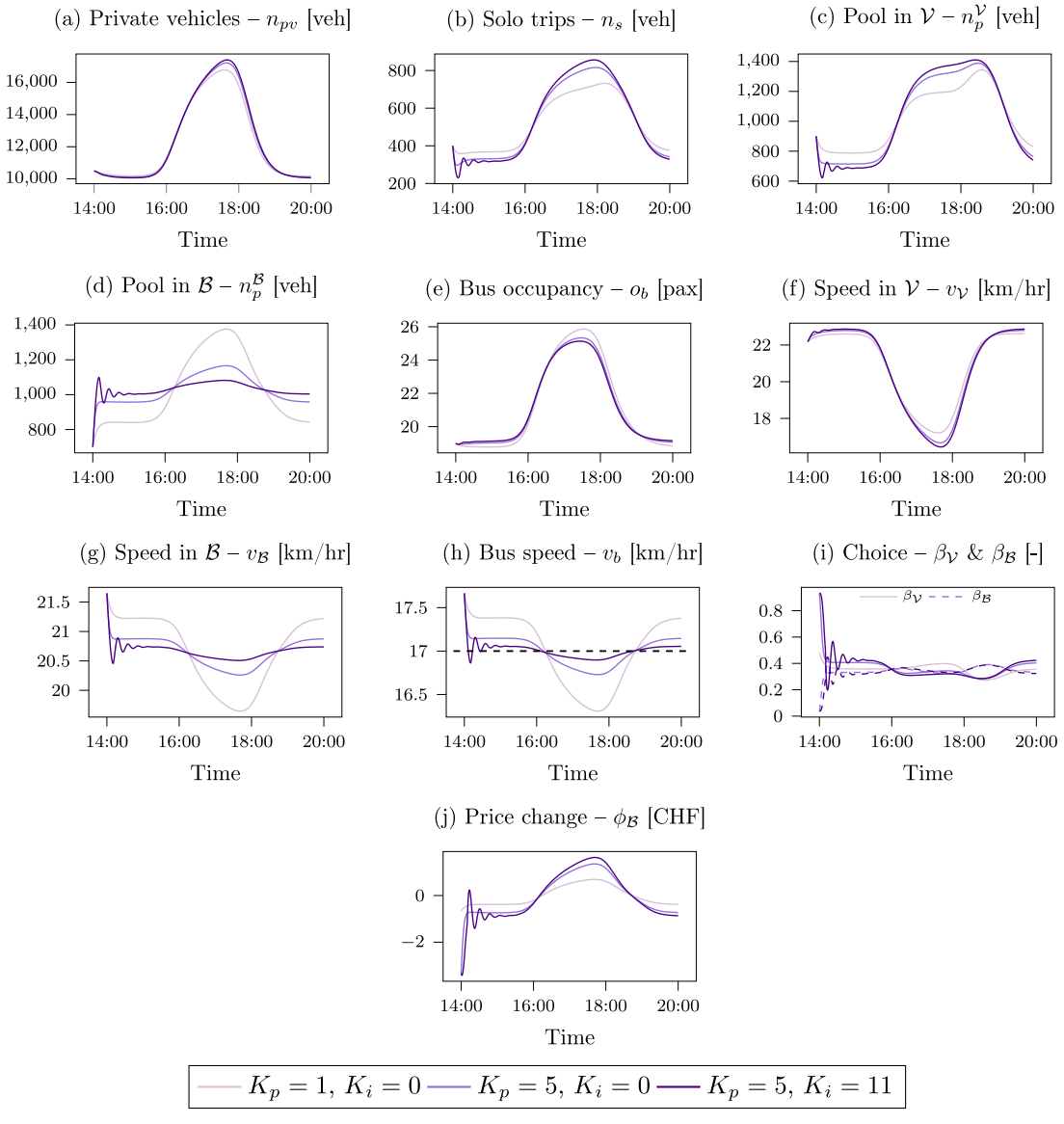

Fig. 6. Time-dependent model variables for the PI framework (Scenario 6) without abandonment for (a) private vehicle accumulation, (b) solo trip ride-hailing vehicles, (c) pool trip ride-hailing accumulation in $\mathcal{V}$, (d) pool trip ride-hailing accumulation in $\mathcal{B}$, (e) average bus occupancy, (f) speed in the vehicle network $\mathcal{V}$, (g) vehicle speed in the bus network $\mathcal{B}$, (h) bus speed in the bus network $\mathcal{B}$, (i) fraction of pool trip in $\mathcal{V}$ and $\mathcal{B}$ respectively, and (j) regulatory control fare for pooling in $\mathcal{B}$.

delays does not fully coincide. Nevertheless, the static choice of $\bar{v}_b$ gives some insights into what value is potentially able to minimize delays. Moreover, despite accounting for abandonment in our dynamic framework, we have observed that the PI displays a similar performance compared to the case without abandonment, for which the comparison with the static framework is performed.

However, despite it being convenient from an implementation point of view, the formulation of the PI framework does not allow for explicitly achieving a multi-modal system optimum. Therefore, we broaden our formulation to include all network users by resorting to an MPC framework, and report the results for when the optimization framework runs with one control variable $\phi_{\mathcal{B}}$, and two control variables $\phi_{\mathcal{V}}$ and $\phi_{\mathcal{B}}$.

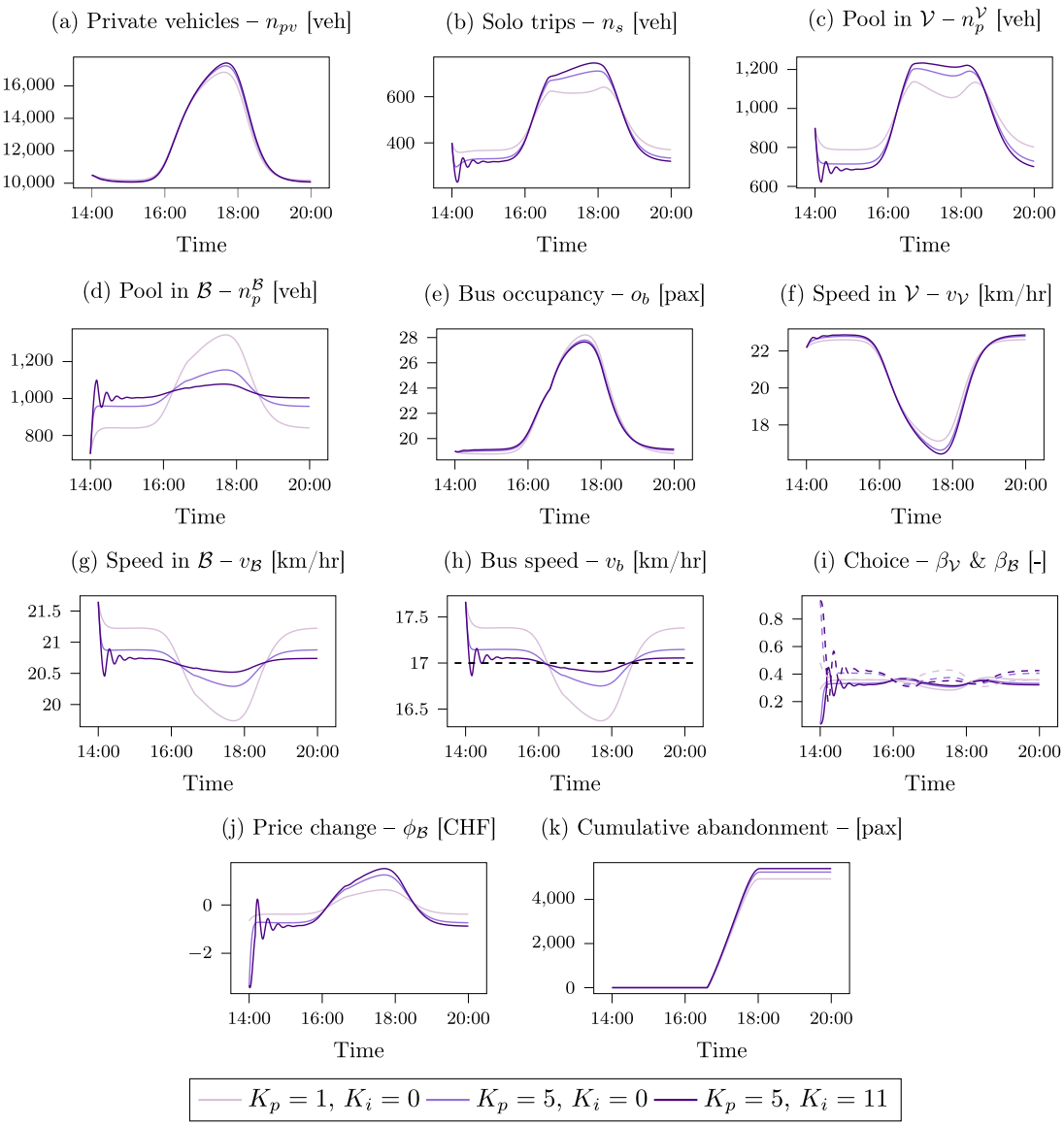

Fig. 7. Time-dependent model variables for the PI framework (Scenario 6) with abandonment for (a) private vehicle accumulation, (b) solo trip ride-hailing vehicles, (c) pool trip ride-hailing accumulation in $\mathcal{V}$, (d) pool trip ride-hailing accumulation in $\mathcal{B}$, (e) average bus occupancy, (f) speed in the vehicle network $\mathcal{V}$, (g) vehicle speed in the bus network $\mathcal{B}$, (h) bus speed in the bus network $\mathcal{B}$, (i) fraction of pool trip in $\mathcal{V}$ and $\mathcal{B}$ respectively, (j) regulatory control fare for pooling in $\mathcal{B}$, and (k) abandoning requests.

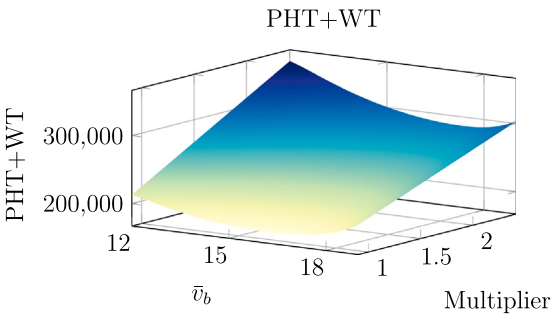

Fig. 8. Objective function values for different set points and different bus demand profiles under the PI control framework without abandonment.

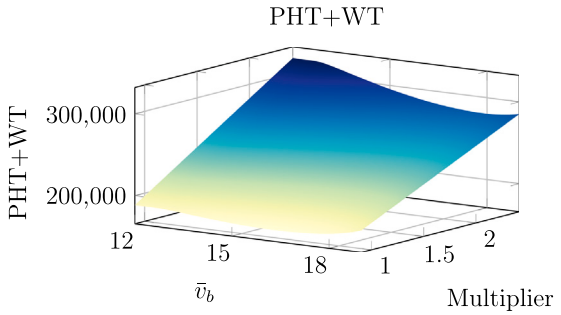

Fig. 9. Objective function values for different set points and different bus demand profiles under the PI control framework with abandonment.

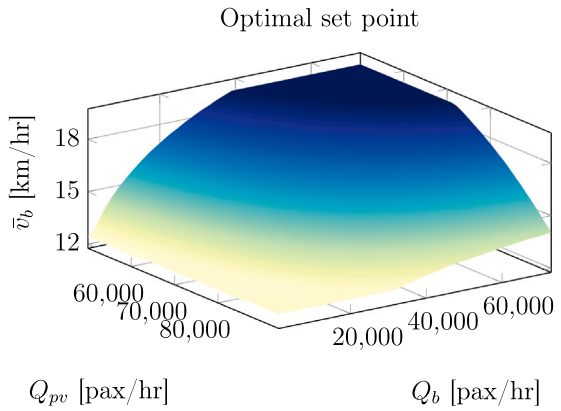

Fig. 10. Optimal PI bus speed set points for different private vehicles and bus demands.

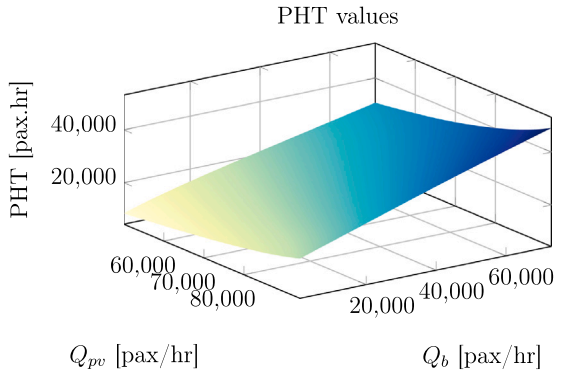

Fig. 11. PHT values for optimal set points.

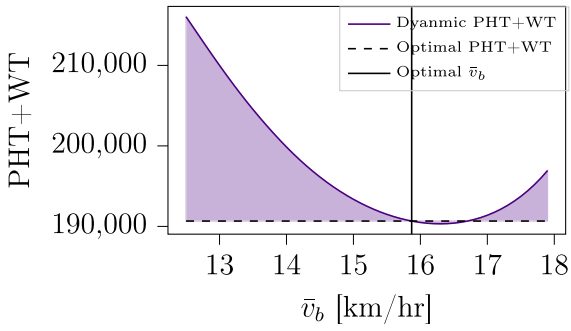

Fig. 12. Difference between the optimal static and dynamic bus speed set points for the PI controller.

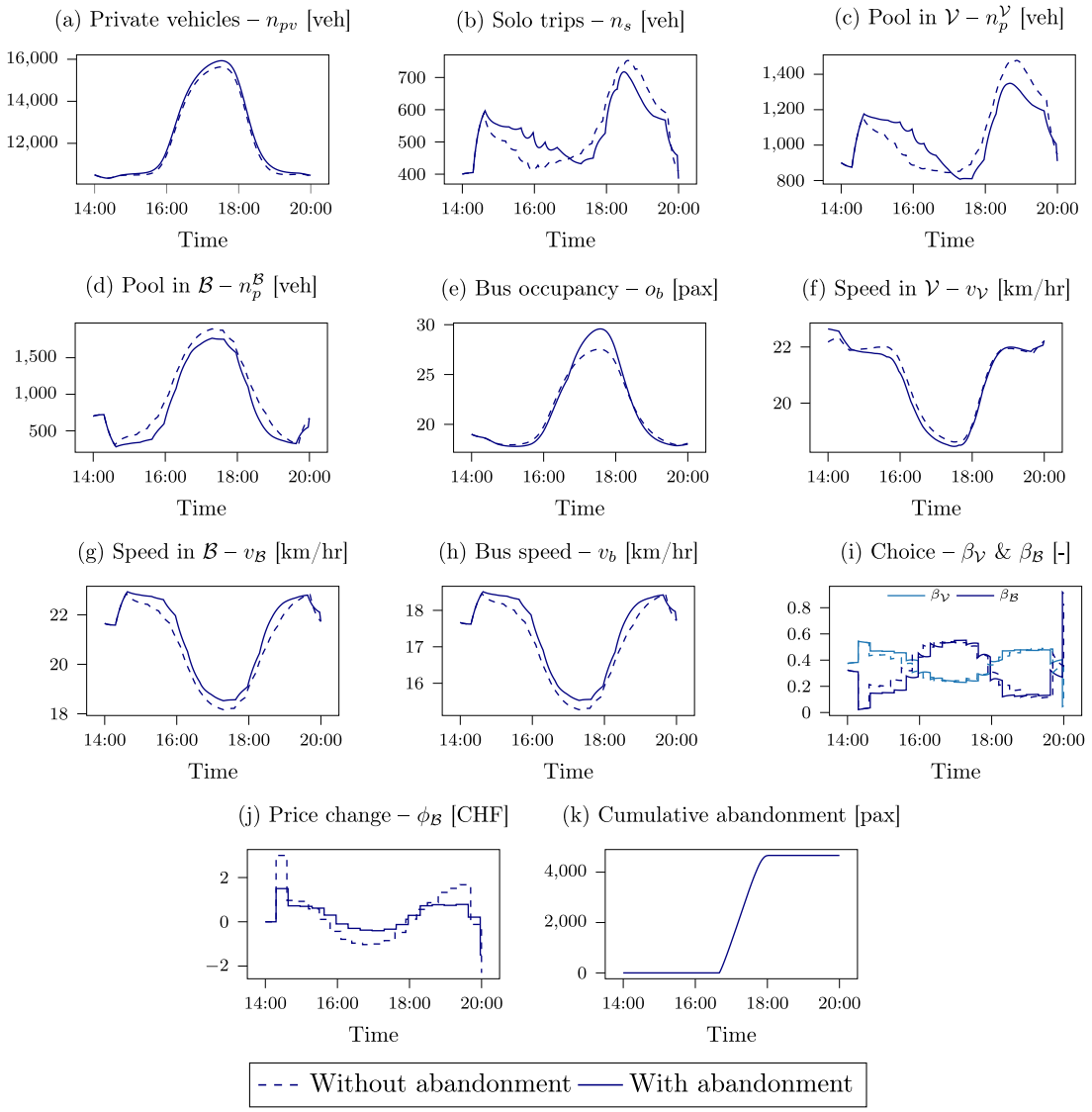

Fig. 13. Time-dependent model variables for Scenario 7 of the MPC framework without and with abandonment for (a) private vehicle accumulation, (b) solo trip ride-hailing vehicles, (c) pool trip ride-hailing accumulation in $\mathcal{V}$, (d) pool trip ride-hailing accumulation in $\mathcal{B}$, (e) average bus occupancy, (f) speed in the vehicle network $\mathcal{V}$, (g) vehicle speed in the bus network $\mathcal{B}$, (h) bus speed in the bus network $\mathcal{B}$, (i) fraction of pool trip in $\mathcal{V}$ and $\mathcal{B}$ respectively, and (j) regulatory control fare for pooling in $\mathcal{B}$, and (k) cumulative abandoning requests.

5.4. MPC framework

We compute the solutions of the MPC problem (20) by utilizing the *JuMP*[1] package in Julia with the *Ipopt*[2] solver. To understand how the MPC framework influences the results, we display the variations of the main simulation variables in Fig. 13 for both the cases without and with abandonment where only the pricing in the bus network $\phi_{\mathcal{B}}$ is controlled. The two implementations show that in the case with abandonment, the difference between the MPC and the plant dynamics originates from the error in the estimation of the number of waiting ride-hailing requests, because request abandonment is solely accounted for within the plant dynamics but not the MPC dynamics. For each optimization run, we consider a prediction horizon of $\tilde{N} = 650$ time steps, i.e., we assume that the optimizer is aware of the private vehicle, bus, and ride-hailing demands for the upcoming 650τ duration. However, after every update of $\tilde{T} = 200$ time steps, we reinitialize the MPC dynamics with the actual state variables for the model dynamics with

[1] https://jump.dev/JuMP.jl/stable/
[2] https://coin-or.github.io/Ipopt/

Table 6
Variation of the MPC objective function value with the prediction horizon.

Prediction horizon [min]	20	25	30	60	100	200	300
PHT+WT [pax.h]	185 104	185 432	184 642	184 755	184 535	184513	184 512

abandonment, to bridge the gap between the MPC predictions and the actual network dynamics. For the case without abandonment, the plant and MPC dynamics are identical, and the prediction horizon value $\tilde{N}$ is therefore set to 3600 time steps in (20) which is the entire simulation duration, thus we assume a perfect knowledge of the future demand. In short, in this case, the abandonment is set to 0 and we solve the MPC in one go. Note that for these simulations, we set $\xi_{\min}$ and $\xi_{\max}$ to be equal to e^{-3} and e^{3} respectively. Moreover, for practical reasons, we only allow the control variable to be updated every 180 time steps such that $N_u = 180$.

As opposed to the PI controller, the MPC shows a more demand-responsive behavior where, during off-peak periods, the speed in the vehicle network in Fig. 13f is lower than that in the bus network observed in Fig. 13g. This substantiates the relatively high values of ϕ_B as the use of the bus network is privileged in this case. During the on-peak period, however, ϕ_B becomes negative, implying that ride-hailing users in the vehicle networks are causing significant delays compared to bus users. Ecouraging users to pool their rides in bus lanes is therefore necessary to reduce the total delays for multi-modal users.

Having described the details of the MPC implementation, we compare the results of the different MPC scenarios in Tables 2 and 4, and show that when our regulatory prices dictate the amount of pooling in both the vehicle and the bus network (Scenario 9), we get the lowest objective function value for the case without abandonment. This is, however, not true for the case with abandonment in Table 4, despite the value of abandonment being the lowest for this specific scenario. The reason for that is that the plant and MPC dynamics are not similar in the case with abandonment, and therefore having controlling two variables instead of one does not necessarily result in lower PHT and WT values. Note that for the no abandonment case, even if we only consider one control variable ϕ_B, we already observe some improvements compared to the no control scenario. In fact, the results in Table 2 reveal that, in reference to the scenario with no control and no pooling in bus lanes, the MPC framework allows for 8% improvement compared to the PI implementation.

For the scenarios with the soft constraint on the bus speed, i.e., Scenarios 8 and 10, the objective function for the MPC becomes

$$\sum_{j=k}^{k+\tilde{N}} \mathrm{PHT}(j) + \mathrm{WT}(j) + \lambda \max\left(0, \underline{v}_b - v_b(j)\right) \tag{21}$$

where $\lambda = 10000$ and $\underline{v}_b$ being equal to 17 km/h. Clearly, when a soft constraint is introduced, the PHT and WT values increase because the private vehicle and ride-hailing delays, which we denote by PHT_{pv} and PHT_{rs}, also increase as reported in Table 3 compared to Scenario 7 with no soft constraints imposed on the bus speed. However, when we numerically compute the bus delays PHT_b and the constraint violations of the minimum bus speed, we guarantee the public transit performance is minimally impacted. The results for the no abandonment control case were added as a sanity check, because by restricting the degrees of freedom of the optimizer, we worsen the value of the objective function between Scenarios 7 and 8, and 9 and 10 in Table 2. Moreover, when the model used in the MPC does not match the real dynamics in the plant, there is no guarantee that the MPC will improve the performance compared to the PI, as can be observed in Table 4 where abandonment is included in the plant.

Finally, to further investigate how the output under the MPC framework changes with different prediction horizons, we display in Table 6 the MPC objective function value for different prediction horizon values $\tilde{N}$ expressed in minutes for the case where a single variable, ξ_B is controlled, and abandonment is only considered in the plant. We note that when $\tilde{N}$ is greater than 600, or equivalently, when the prediction horizon is greater than 60 min, a further increase in the prediction horizon will not improve the objective function value significantly.

6. Conclusions

In this work, we develop an aggregate dynamic for a multi-modal network with private vehicles, ride-hailing services, and public transportation. Our modeling approach aims at evaluating an occupancy-dependent vehicle allocation policy where a fraction of the pool ride-hailing users choose to utilize dedicated bus lanes. Despite ameliorating the total user delays and the ride-hailing service levels, our allocation strategy is not capable of influencing by itself the ride-hailing user choices in a way that aligns with the sharing split that minimizes these overall multi-modal user delays. The need for a more elaborate pricing scheme to steer the user's choice towards a more convenient solution for multi-modal users is therefore substantiated. Consequently, we build both a PI and an MPC control framework, and analyze what the additional pooling discount or fare should be that improves the overall travel times for all network users, and the waiting time for ride-hailing users in particular. Our results show that pricing indeed influences the preferences of ride-hailing requests in a manner that reduces the PHT for all modes. We performed the same analysis for the cases with and without request abandonment, and demonstrated how the complexity of our abandonment function is circumvented in the MPC solution.

In future work, we plan to give further attention to demand-dependent trip detours. This is because the detour distance for a pool trip is lower when more passengers opt for pooling. Therefore, the efficiency of our proposed policy becomes more accentuated if this factor is accounted for in the modeling formulation. Moreover, we have so far considered pool trips with only two passengers sharing their rides. A potential direction for this framework is to extend this work to incorporate a high-capacity on-demand micro-transit

service utilizing the bus network along with buses, and observe the additional occupancy-dependent improvements that we achieve if these services progressively gain more and more momentum. Furthermore, the macroscopic traffic model formulated for the two networks under consideration overlooks the potential interactions between the different modes utilizing the separate parts of the network. This interaction requires refining the network-based traffic model to include for instance the likelihood that pool vehicles in bus lanes utilize the vehicle network to overtake buses, especially in roads where the space is constrained. Finally, a crucial point to investigate is driver's compliance to the matching decisions by the ride-hailing platform, and how this could be integrated within the proposed modeling framework.

CRediT authorship contribution statement

Lynn Fayed: Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Gustav Nilsson:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Conceptualization. **Nikolas Geroliminis:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Conceptualization.

Acknowledgments

This work was supported by the Swiss National Science Foundation under NCCR Automation, grant agreement 51NF40_180545.

References

Alisoltani, N., Leclercq, L., Zargayouna, M., 2021. Can dynamic ride-sharing reduce traffic congestion? Transp. Res. B: Methodological http://dx.doi.org/10.1016/j.trb.2021.01.004.

Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., Rus, D., 2017. On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. Proc. Natl. Acad. Sci. 114, 201611675. http://dx.doi.org/10.1073/pnas.1611675114.

Beojone, C.V., Geroliminis, N., 2021. On the inefficiency of ride-sourcing services towards urban congestion. Transp. Res. C 124, 102890. http://dx.doi.org/10.1016/j.trc.2020.102890.

Bimpikis, K., Candogan, O., Saban, D., 2019. Spatial pricing in ride-sharing networks. Oper. Res. 67 (3), 744–769. http://dx.doi.org/10.1287/opre.2018.1800, arXiv:10.1287/opre.2018.1800.

Cachon, G.P., Daniels, K.M., Lobel, R., 2017. The role of surge pricing on a service platform with self-scheduling capacity. Manuf. Serv. Oper. Manage. 19 (3), 368–384. http://dx.doi.org/10.1287/msom.2017.0618.

Castillo, J.C., Knoepfle, D., Weyl, G., 2017. Surge pricing solves the wild goose chase. In: Proceedings of the 2017 ACM Conference on Economics and Computation. EC '17, Association for Computing Machinery, New York, NY, USA, pp. 241–242. http://dx.doi.org/10.1145/3033274.3085098.

Cohen, M.C., Jacquillat, A., Ratzon, A., Sasson, R., 2022. The impact of high-occupancy vehicle lanes on carpooling. Transp. Res. A: Policy and Practice 165, 186–206. http://dx.doi.org/10.1016/j.tra.2022.08.021.

Cramer, J., Krueger, A.B., 2016. Disruptive change in the taxi business: The case of uber. Amer. Econ. Rev. 106 (5), 177–182. http://dx.doi.org/10.1257/aer.p20161002.

de Souza Silva, L.A., de Andrade, M.O., Alves Maia, M.L., 2018. How does the ride-hailing systems demand affect individual transport regulation? Res. Transp. Econom. 69, 600–606. http://dx.doi.org/10.1016/j.retrec.2018.06.010.

Diao, M., Kong, H., Zhao, J., 2021. Impacts of transportation network companies on urban mobility. Nat. Sustain. 4, 1–7. http://dx.doi.org/10.1038/s41893-020-00678-z.

Erhardt, G., Roy, S., Cooper, D., Sana, B., Chen, M., Castiglione, J., 2019. Do transportation network companies decrease or increase congestion? Sci. Adv. 5, eaau2670. http://dx.doi.org/10.1126/sciadv.aau2670.

Fayed, L., Nilsson, G., Geroliminis, N., 2023a. A macroscopic modelling framework for the dynamic pricing of pool ride-splitting vehicles in bus lanes. In: 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC). pp. 1657–1662. http://dx.doi.org/10.1109/ITSC57777.2023.10421875.

Fayed, L., Nilsson, G., Geroliminis, N., 2023b. On the utilization of dedicated bus lanes for pooled ride-hailing services. Transp. Res. B 169, 29–52. http://dx.doi.org/10.1016/j.trb.2023.01.005.

Geroliminis, N., Daganzo, C.F., 2008. Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. Transp. Res. B 42 (9), 759–770. http://dx.doi.org/10.1016/j.trb.2008.02.002.

Geroliminis, N., Zheng, N., Ampountolas, K., 2014. A three-dimensional macroscopic fundamental diagram for mixed bi-modal urban networks. Transp. Res. C 42, 168–181. http://dx.doi.org/10.1016/j.trc.2014.03.004.

Guo, X., Wang, Q., Zhao, J., 2022. Data-driven vehicle rebalancing with predictive prescriptions in the ride-hailing system. IEEE Open J. Intell. Transp. Syst. 3, 251–266. http://dx.doi.org/10.1109/OJITS.2022.3163180.

Hall, J.D., Palsson, C., Price, J., 2018. Is Uber a substitute or complement for public transit? J. Urban Econom. 108, 36–50. http://dx.doi.org/10.1016/j.jue.2018.09.003.

Jiao, G., Ramezani, M., 2022. Incentivizing shared rides in e-hailing markets: Dynamic discounting. Transp. Res. Part C: Emerg. Technol. http://dx.doi.org/10.1016/j.trc.2022.103879.

Jung, J., Jayakrishnan, R., Park, J.Y., 2016. Dynamic shared-taxi dispatch algorithm with hybrid-simulated annealing. Comput.-Aided Civ. Infrastruct. Eng. 31 (4), 275–291. http://dx.doi.org/10.1111/mice.12157, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/mice.12157.

Ke, J., Yang, H., Li, X., Wang, H., Ye, J., 2020. Pricing and equilibrium in on-demand ride-pooling markets. Transp. Res. B 139, 411–431. http://dx.doi.org/10.1016/j.trb.2020.07.001.

Ke, J., Zheng, Z., Yang, H., Ye, J., 2021a. Data-driven analysis on matching probability, routing distance and detour distance in ride-pooling services. Transp. Res. C 124, 102922. http://dx.doi.org/10.1016/j.trc.2020.102922.

Ke, J., Zhu, Z., Yang, H., He, Q., 2021b. Equilibrium analyses and operational designs of a coupled market with substitutive and complementary ride-sourcing services to public transits. Transp. Res. Part E: Logist. Transp. Rev. 148, 102236. http://dx.doi.org/10.1016/j.tre.2021.102236.

Lamotte, R., de Palma, A., Geroliminis, N., 2017. On the use of reservation-based autonomous vehicles for demand management. Transp. Res. B 99, 205–227. http://dx.doi.org/10.1016/j.trb.2017.01.003.

Li, S., Qin, J., Yang, H., Poolla, K., Varaiya, P., 2020. Off-street parking for TNC vehicles to reduce cruising traffic. In: 2020 59th IEEE Conference on Decision and Control (CDC). pp. 2585–2590.

Loder, A., e, L., Menendez, M., Axhausen, K.W., 2017. Empirics of multi-modal traffic networks – using the 3D macroscopic fundamental diagram. Transp. Res. C 82, 88–101. http://dx.doi.org/10.1016/j.trc.2017.06.009.

Ma, M., Chen, Y., Liu, W., Waller, S.T., 2023. An economic analysis of a multi-modal transportation system with ride-sourcing services and multi-class users. Transp. Policy 140, 1–17. http://dx.doi.org/10.1016/j.tranpol.2023.06.008.

Mo, B., Cao, Z., Zhang, H., Shen, Y., Zhao, J., 2021. Competition between shared autonomous vehicles and public transit: A case study in Singapore. Transp. Res. C 127, http://dx.doi.org/10.1016/j.trc.2021.103058.

Ni, W., Cassidy, M., 2019. City-wide traffic control: Modeling impacts of cordon queues. Transp. Res. C 113, http://dx.doi.org/10.1016/j.trc.2019.04.024.

Nourinejad, M., Ramezani, M., 2020. Ride-sourcing modeling and pricing in non-equilibrium two-sided markets. Transp. Res. B 132, 340–357. http://dx.doi.org/10.1016/j.trb.2019.05.019, 23rd International Symposium on Transportation and Traffic Theory (ISTTT 23).

Ramezani, M., Valadkhani, A.H., 2023. Dynamic ride-sourcing systems for city-scale networks - Part I: Matching design and model formulation and validation. Transp. Res. C 152, 104158. http://dx.doi.org/10.1016/j.trc.2023.104158.

Shaheen, S., Bansal, A., Chan, N., Cohen, A., 2017. Mobility and the sharing economy: Industry developments and early understanding of impacts. Low Carbon Mobility for Future Cities http://dx.doi.org/10.1049/PBTR006E_CH10.

Shaheen, S., Cohen, A., 2019. Shared ride services in North America: Definitions, impacts, and the future of pooling. Transp. Rev. 39 (4), 427–442. http://dx.doi.org/10.1080/01441647.2018.1497728, arXiv:10.1080/01441647.2018.1497728.

Sirmatel, I.I., Tsitsokas, D., Kouvelas, A., Geroliminis, N., 2021. Modeling, estimation, and control in large-scale urban road networks with remaining travel distance dynamics. Transp. Res. C http://dx.doi.org/10.1016/j.trc.2021.103157.

Soza-Parra, J., Kucharski, R., Cats, O., 2023. The shareability potential of ride-pooling under alternative spatial demand patterns. Transportmetrica A Transp. Sci. http://dx.doi.org/10.1080/23249935.2022.2140022.

Tachet, R., Sagarra, O., Santi, P., Resta, G., Szell, M., Strogatz, S., Ratti, C., 2016. Scaling law of urban ride sharing. Sci. Rep. http://dx.doi.org/10.1038/srep42868.

Tirachini, A., 2020. Ride-hailing, travel behaviour and sustainable mobility: An international review. Transportation 47, http://dx.doi.org/10.1007/s11116-019-10070-2.

Tirachini, A., Gómez-Lobo, A., 2019. Does ride-hailing increase or decrease vehicle kilometers traveled (VKT)? A simulation approach for Santiago de Chile. Int. J. Sustain. Transp. 14, 1–18. http://dx.doi.org/10.1080/15568318.2018.1539146.

Toledo, T., Mansour, O., Haddad, J., 2017. Optimal dynamic tolls for managed lanes. Transp. Res. Rec.: J. Transp. Res. Board 2606, 28–37. http://dx.doi.org/10.3141/2606-04.

Tsitsokas, D., Kouvelas, A., Geroliminis, N., 2021. Modeling and optimization of dedicated bus lanes space allocation in large networks with dynamic congestion. Transp. Res. C 127, 103082. http://dx.doi.org/10.1016/j.trc.2021.103082.

Vignon, D., Yin, Y., Ke, J., 2023. Regulating the ride-hailing market in the age of uberization. Transp. Res. Part E: Logist. Transp. Rev. 169, 102969. http://dx.doi.org/10.1016/j.tre.2022.102969.

Wang, J., Wang, X., Yang, S., Yang, H., Zhang, X., Gao, Z., 2021. Predicting the matching probability and the expected ride/shared distance for each dynamic ridepooling order: A mathematical modeling approach. Transp. Res. B 154, 125–146. http://dx.doi.org/10.1016/j.trb.2021.10.005.

Xu, Z., Yin, Y., Ye, J., 2020. On the supply curve of ride-hailing systems. Transp. Res. B 132, 29–43. http://dx.doi.org/10.1016/j.trb.2019.02.011.

Xu, Z., Yin, Y., Zha, L., 2017. Optimal parking provision for ride-sourcing services. Transp. Res. B 105, 559–578. http://dx.doi.org/10.1016/j.trb.2017.10.003.

Yang, Y., Ramezani, M., 2023. A learning method for real-time repositioning in E-hailing services. IEEE Trans. Intell. Transp. Syst. http://dx.doi.org/10.1109/TITS.2022.3219381.

Yu, J.J., Tang, C.S., Max Shen, Z.-J., Chen, X.M., 2020. A balancing act of regulating on-demand ride services. Manage. Sci. 66 (7), 2975–2992. http://dx.doi.org/10.1287/mnsc.2019.3351, arXiv:10.1287/mnsc.2019.3351.

Zha, L., Yin, Y., Xu, Z., 2018. Geometric matching and spatial pricing in ride-sourcing markets. Transp. Res. C 92, 58–75. http://dx.doi.org/10.1016/j.trc.2018.04.015.

Zha, L., Yin, Y., Yang, H., 2016. Economic analysis of ride-sourcing markets. Transp. Res. C 71, 249–266. http://dx.doi.org/10.1016/j.trc.2016.07.010.

Zhang, K., Nie, Y.M., 2021. To pool or not to pool: Equilibrium, pricing and regulation. Transp. Res. B 151, 59–90. http://dx.doi.org/10.1016/j.trb.2021.07.001.

Zhu, Z., Xu, A., He, Q.-C., Yang, H., 2021. Competition between the transportation network company and the government with subsidies to public transit riders. Transp. Res. Part E: Logist. Transp. Rev. 152, 102426. http://dx.doi.org/10.1016/j.tre.2021.102426.