http://dx.doi.org/10.1016/j.trc.2021.102987

Integrated optimal control strategies for freeway traffic mixed with connected automated vehicles: A model-based reinforcement learning approach

Tianlu Pan [a], Renzhong Guo [a], William H.K. Lam [b], Renxin Zhong [c,*], Weixi Wang [a], Biao He [a,*]

[a] *Guangdong Key Laboratory of Urban Informatics, School of Architecture and Urban Planning, Shenzhen University, Shenzhen, China.*
[b] *Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China.*
[c] *School of Intelligent Systems Engineering, Sun Yat-Sen University, Guangzhou, China.*

Abstract

This paper proposes an integrated freeway traffic flow control framework that aims to minimize the total travel cost, improve greenness and safety for freeway traffic mixed with connected automated vehicles (CAVs) and regular human-piloted vehicles (RHVs). The proposed framework devises an integrated action of several control strategies such as ramp metering, lane changing control (LCC) for CAVs and lane changing recommendation (LCR) for RHVs, variable speed limit control (VSLC) for CAVs and variable speed limit recommendation (VSLR) for RHVs with minimum safety gap control measures for lane changing and merging maneuvers. The CAVs are assumed to follow the system control instructions fully and immediately. In contrast, the RHVs would make decisions in response to the recommendations disseminated and also the behaviors of CAVs. The compliance rate of drivers to the LCR is captured by the underlying traffic flow model. A set of constraints is imposed to restrict VSLC/VSLR and LCC/LCR measures from changing too frequently or too sharply on both temporal and spatial dimensions to avoid excessive nuisance to passengers and traffic flow instability. A reinforcement learning based solution algorithm is proposed. First, a control parameterization technique is adopted to reduce the dimension of the original optimal control problem to increase computational efficiency. Then, a gradient-free Cross-Entropy-Method based algorithm is used to search the optimal parameters to circumvent the non-differentiability of the traffic flow model. The feasibility and effectiveness of the proposed framework are illustrated via numerical examples for a variety of penetration rates of CAVs under various traffic conditions. A sensitivity analysis is conducted to demonstrate the impacts of several important parameters such as the reaction time of the CAVs. It is found that the integrated control strategy can reduce the total travel cost by reducing the lane changing maneuvers and vehicles queuing at the bottleneck meanwhile smooth the traffic flow and suppress the adverse impact of shockwaves. The effect of ramp metering is not significant when the penetration rate of CAVs is high enough. Speed harmonization (with minimum gap control) in conjunction with LCC/LCR would be a better integrated control strategy under high penetration rate of CAVs.

Keywords: Integrated traffic control; Vehicle automation and communication system; Multiclass multilane cell transmission model; Connected automated vehicle; Penetration rate.

1. Introduction

Considerable effort has been dedicated to develop vehicle automation and communication systems (VACS) in the last decades. The VACS are expected to revolutionize the features and capabilities of vehicles via connecting

*Corresponding authors.
E-mail addresses: glorious9009@gmail.com (Tianlu Pan)., guorz2013@qq.com (Renzhong Guo)., william.lam@polyu.edu.hk (William H.K. Lam)., zhrenxin@mail.sysu.edu.cn (Renxin Zhong)., 359912469@qq.com (Weixi Wang)., 1688698@qq.com (Biao He).

vehicles with infrastructures and enabling autonomous driving in the next decades. Recent research has indicated that the penetration of connected automated vehicles (CAVs) (e.g., automated vehicles (AV) connected via VACS) can improve traffic safety and the traffic efficiency while reducing traffic induced emission through traffic control (Diakaki et al., 2015; Zhu and Ukkusuri, 2015; Roncoli et al., 2015a,b, 2016; Stern et al., 2017). However, before a high penetration of CAVs, the market will be still dominated by regular human-piloted vehicles (RHVs) in the coming decades (Levin and Boyles, 2016a,b). As a consequence, the road traffic will be mixed with the CAVs and RHVs in the near future.

The CAVs can reduce reaction time so as to improve traffic flow efficiency via smaller inter-vehicle headway and the connected vehicle platooning can directly change the traffic flow characteristics with adaptive cruising control system (ACC) and cooperative adaptive cruising control system (CACC) schemes (Levin and Boyles, 2016a; Jia and Ngoduy, 2016a,b; Gao et al., 2017; Zhou et al., 2017; Zhu and Ukkusuri, 2018; Gao et al., 2019). Given the expected penetration rate of AVs, Mohajerpoor and Ramezani (2019) derived the headway distribution of a traffic stream with mixed AVs and RHVs from a microscopic traffic description. The CAVs also give the possibility to implement control schemes such as individual vehicle speed and lane selection advice that are not available with RHVs (Han et al., 2017; Tian et al., 2019). However, rather limited research effort has been dedicated to addressing the implications of the emerging CAVs on the flow characteristics of traffic mixed with CAVs and RHVs, as well as their potential exploitation for improving traffic flow operations.

Several concepts concerning traffic control systems with VACS have been elaborated in either non-cooperative or cooperative manner. Most of these concepts pertain to a multi-layer structure to tackle the huge problem complexity as inspired by Varaiya (1993), see e.g., Diakaki et al. (2015); Roncoli et al. (2015a,b). For instance, the cruise system can automatically adjust the speed of equipped vehicles to follow the preceding vehicles safely and comfortably. By cooperating the performance of a vehicle platoon with VACS, the platooning vehicles can respond to disruptions in traffic flow more smoothly while maintaining smaller inter-vehicle gaps compared with RHVs. Another category that has received considerable attention recently is the speed regulation systems. For instance, the extension of the conventional variable speed limit (VSL) system that employs infrastructure-to-vehicle (I2V) communication can disseminate speed limit information determined according to the current speed and position of each vehicle. The lane-changing assistance systems that assist the lane changing and merging maneuvers also form an important category of VACS (Diakaki et al., 2015). It is found that weaving sections of freeways usually admit capacity reductions caused by lane-changing (LC) maneuvers, which have significant impacts on traffic safety. As deemed by Diakaki et al. (2015), to improve freeway safety and efficiency with the aid from VACS, automatic controls regarding speed, headway, lane assignment and lane-change/merge maneuvers of vehicles should be devised.

For congested traffic, any disturbance such as vehicle merging and lane changing can create a shockwave, which may eventually cause traffic breakdown. Microscopic models would be better choices for describing these phenomena. At a microscopic level, car-following controls are developed to optimize local situations of equipped vehicles for safety performance using vehicle propulsion and brake systems. Wang et al. (2016a,b) extended the VSL control strategies wherein the CAVs are used to actuate the traffic flow. Using several single-lane ring road experiments, Stern et al. (2017) demonstrated that the mixed-autonomy can reduce over 40% fuel consumption by inserting a CAV in the traffic to dampen the ring instability. He et al. (2017) proposed a jam absorbing strategy, which advocates a “slow-in, fast-out" driving strategy, based on a simplified Newell’s car following model. Jia et al. (2019) proposed a multiclass microscopic model to describe traffic dynamics mixed with CAVs and RHVs with a consensus-based control algorithm. The disadvantage of microscopic approaches is the scalability that the computational burden increases with the number of cooperative vehicles in the platoon. Moreover, the lane-changing maneuvers are not considered.

At the macroscopic level, ramp metering aims to improve freeway traffic conditions by regulating ramp flows to the freeway mainstream and flows at freeway-to-freeway intersections at the price of introducing short delays at the on-ramps and freeway-to-freeway intersections (Papageorgiou and Kotsialos, 2002). It is found that the ramp metering is useful only when traffic is not too light (otherwise ramp metering is not needed) or not too dense (otherwise traffic breakdown will happen anyway) (Papageorgiou and Kotsialos, 2002; Hegyi et al., 2005a,b). On the other hand, the VSL control is usually implemented upstream to a bottleneck yet to be activated to reduce mainstream flow arriving the bottleneck to retard the bottleneck activation by dissimilating the VSL information via the variable message sign (VMS) gantries (Hegyi et al., 2005a,b; Carlson et al., 2010, 2011; Li et al., 2014, 2016). It is envisaged that the VSL can alleviate freeway traffic congestion through homogenization by reducing the fluctuations in traffic variables. Conventionally, homogenization VSL should be applied at flow rates slightly below capacity (around 15-

20%) by imposing speed limits around the critical speed (around 70-90 km/h) (Papageorgiou et al., 2008). Another popular approach uses VSL as a mainline metering in coordination with ramp metering (RM) to avoid the activation of bottlenecks and the harmful capacity drop by imposing low speed limits upstream of the critical bottlenecks (even down to 10 km/h in Müller et al. (2015) or around 20 km/h in Carlson et al. (2010, 2011)). For mixed freeway traffic, Han et al. (2017) developed VSL strategies to improve bottleneck discharge rates and reduce system delays using CAVs under various penetration rates wherein the traffic evolves according to the kinematic wave theory. It is found that, in contrast to VMS-oriented strategies, the CAV-based strategies can effectively impose dynamic control over continuous time and space.

Taking advantage of CAV technology, a popular extension of the VSL for CAVs is the speed harmonization that aims to reduce temporal and spatial variations of traffic speed, to increase safety and mobility of the transportation systems. Malikopoulos et al. (2019) developed an optimal control-based speed harmonization policy for CAV. It is found that the speed harmonization can reduce 19-22% fuel consumption for each vehicle compared to the scenario with RHVs. Khondaker and Kattan (2015) devised an I2V-based VSL system, where CAVs with CACC functions can receive downstream information much in advance. Based on a sensitivity analysis, it was found that the I2V-based VSL system can improve the uncontrolled scenario by 20% in travel time reductions, and 5-16% in energy consumption reductions, respectively.

As claimed in Soriguera et al. (2017), it lacks empirical evidence to prove the effectiveness of the low speed limit while the traffic flow models used for numerical variable speed limit control (VSLC) studies have never been validated for (very) low speed limits. Mining the empirical data in the B-23 freeway in Barcelona (Spain), Soriguera et al. (2017) found that low-speed limit strategies would fail to restrict the mainline flow. Surprisingly, for moderate demand, results showed that low speed limits could increase the speed differences across lanes which in turn increases the LC rate and worsen the existing congestion problems. The homogenization of traffic and reduction in LC maneuvers via low-speed limits might not be successful in this sense. Note that any strategy that could reduce LC maneuvers could increase flow rates. To explore the VSL homogenization effect while incorporating the LC maneuvers, Zhang and Ioannou (2017) found that the VSL control will have limited or no effect in improving travel time under heavy traffic conditions because there is no way for the VSL control technique to eliminate the capacity drop caused by the low-speed lane-changing traffic. It was concluded that pure VSL control cannot prevent traffic breakdowns in a satisfactory manner under over-saturated traffic conditions. Without combining lane changing control (LCC), once the low speed LC maneuvers bring down the speed of traffic in adjacent lanes, the capacity drop takes place. These findings shed light on the design of speed harmonization schemes for controlling the speed and lane change.

Cooperative merging of CAVs at freeway on-ramps has been extensively studied. A widely adopted approach is to use the concept of virtual vehicles that a virtual vehicle is mapped onto a CACC system of the freeway mainline before the actual merging maneuver. Although the merging maneuver usually involves a lateral motion control of CAVs, i.e., lane changing. The literature mainly focuses on cooperative longitudinal motion control of CAVs to regulate vehicle longitudinal gaps before they reach the conflict zone using VACS (Wang et al., 2020). However, the lateral control problem or the case of traffic mixed RHVs and CAVs is relatively lacking. Most of the existing traffic control strategies for traffic flow mixed with CAVs do not consider the LCC. Soriguera et al. (2017) encouraged future effort first to develop better traffic models to emulate the effects regarding VSL and LC maneuvers, and then to overcome the above limitations of VSL by incorporating the LCC. To consider the LCC, lane-level models that can capture vehicle LC maneuvers are preferable (Pan et al., 2016; Zhang and Ioannou, 2017; Pan et al., 2021). At the lane-level, the VSL can be used to reduce speed differences among vehicles traveling in the same lane and/or adjacent lanes. As the speed difference is the major incentive for discretionary LC, this reduction in speed differences can synchronize drivers' behavior and suppress LC maneuvers, thereby decreases the probability of traffic collision. Tian et al. (2019) showed that up to 8% total travel times of connected vehicles can be reduced by lane selection assistance application.

To sum up, a single traffic control scheme may not be sufficient for complex traffic scenarios. Integration of traffic control strategies such as LCC, VSLC, and minimum safety gap control is necessary for improving traffic efficiency and safety (Talebpour and Mahmassani, 2016; Han et al., 2017; Zhang and Ioannou, 2017; Tian et al., 2019). This paper tackles the above challenges by proposing an integrated freeway traffic control framework for traffic mixed with CAVs and RHVs. The main contributions of this paper are:

- The proposed control framework employs an integrated action of a number of control measures, including ramp metering, variable speed limit at the link level, and lane changing control at the lane level. The minimum

safety gap control, implicitly included in the proposed traffic model, is built for the safety of car following, lane changing and merging maneuvers. This integrated control can minimize the total travel cost, improve greenness and safety via reducing the fluctuation of control for freeway traffic mixed with CAVs and RHVs.

- Noting that the drivers would not fully follow the LCR, we capture their responses to the LCR disseminated by the en-route Variable Message Signs and navigation APPs, i.e., the (cumulative) compliance of drivers to the LCR, via a validated exponential like function embedded in the proposed traffic flow model.
- A set of constraints is imposed to restrict VSLC/VSLR and LCC/LCR measures from changing too frequently or too sharply on both temporal and spatial dimensions to avoid excessive nuisance to passengers and traffic flow instability.
- A model-based reinforcement learning solution algorithm is proposed. Due to the complexity of the multiclass multilane traffic flow model, the optimization problem is with non-differentiable functional structure of the control vector and a complicated constraint set that would induce many local optima. Rather than searching the value function iteration, this algorithm parameterizes the control (to reduce the dimension of the original optimal control problem to increase computational efficiency) and searches for the optimal parameters that lead to minimizing the objective function via a gradient-free Cross-Entropy-Method based algorithm to circumvent the non-differentiability of the traffic flow model.

This paper is organized as follows. An optimal control problem is formulated in Section 2. A multiclass multilane traffic model extended from Pan et al. (2021) is introduced in Section 3 to describe flow characteristics of traffic mixed with RHVs and CAVs under various traffic control schemes. A reinforcement learning based gradient-free CEM algorithm is developed to solve the optimal control problem in Section 4. Numerical simulations are conducted in Section 5 to assess the performance of the integrated control and to analyze the impact of penetration rate of CAVs and the congestion level on the control performance. Finally, Section 6 concludes the paper and depicts the future works. Companion materials are presented in the appendix.

2. Optimal control strategies based on multiclass traffic flow modeling

2.1. Settings of the optimal control problem

This section formulates the optimal control problem for freeway traffic with a given penetration rate of CAVs. The control architecture considers a central traffic manager that devises and evaluates the optimal control problem and disseminates the control strategies to both CAVs and RHVs. It is assumed that the traffic manager has complete knowledge of the traffic state, which is estimated by the multiclass multilane traffic flow model extended from Pan et al. (2021). In line with the literature, we first assume that the penetration rate of CAVs is sufficient and that the control can actually be implemented via appropriate actuators.

The objective of the optimal control problem is to improve the traffic efficiency, greenness and safety through an integration of ramp metering, VSLC and LCC for CAVs, the VSLR and LCR for RHVs, while improving safety by enforcing minimum safety gap control. As shown in Figure 1, the vehicle-to-vehicle (V2V) system measures the vehicular states, such as (relative) speed, acceleration, and gap between adjacent vehicles, and transmits the data to feed the controller. The CAVs are assumed to fully follow the control strategies issued by VACS through Vehicle-to-Everything Road-Side-Unit (V2X RSU). On the other hand, the VSLR and LCR are released via the existing infrastructures such as VMS gantries and navigation APPs. In the present work, the VSL is specified by the traffic manager for each segment-lane such that all CAVs traveling on the segment-lane will follow the respective speed or speed limit. As reported in the literature, the average speed of RHVs is roughly equal to the average speed of connected vehicles in a traffic stream (Bekiaris-Liberis et al., 2016; Fountoulakis et al., 2017). Therefore, it will be sufficient to impose the speed limit, even on RHVs, given an adequate penetration of equipped vehicles, e.g., there is at least one CAV within each segment-lane. Therefore, we do not consider the compliance of drivers to VSL control in this paper. The effect of LCR depends on the compliance of drivers as well as the impact of the behavior of CAVs. In particular, the compliance of LCR is assumed to follow an exponential like function, which is embedded in the multiclass multilane traffic flow model and was validated in Pan et al. (2016). Since we concentrate on a mixed traffic scenario that the CAVs and RHVs are randomly arriving at the on-ramps while the RHVs cannot make

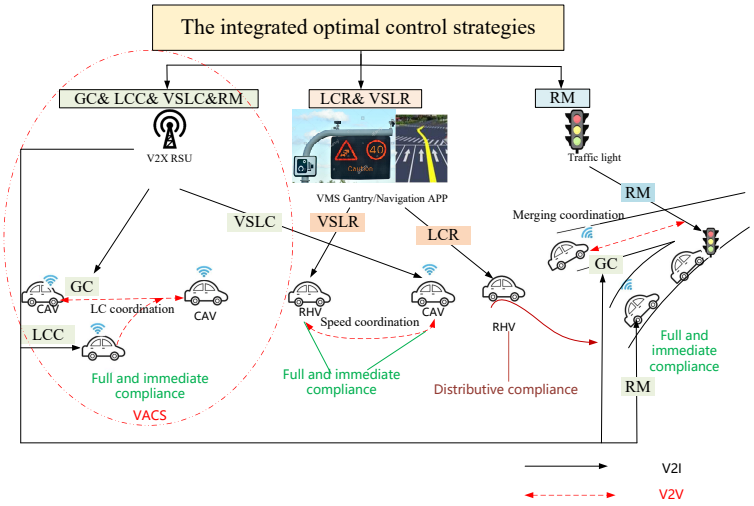


Figure 1: An illustration of the integrated control strategies to CAVs and RHVs

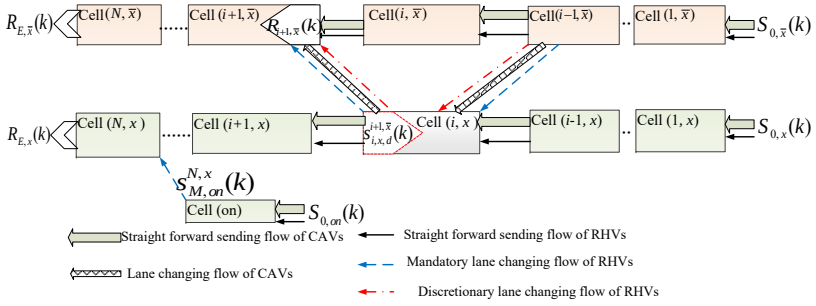


Figure 2: Merging and diverging induced by lane-changing of CAVs and RHVs

a cooperative decision with the CAVs, we adopt the conventional signal-based ramp metering control. **Within the mixed traffic flow, both VSLC/VSLR should avoid too frequent and sharp amplitude for escaping excessive nuisance to drivers and passengers. Meanwhile, the LCR and RM should also avoid frequent fluctuations or unrealistic setting.** Such issues are considered by imposing a set of constrains to the corresponding decision variables and choosing proper control cycles in Subsection 3.2.

2.2. Formulation of the optimal control problem

We consider a dual-lane freeway segment divided into several cell packages, from cell package 1 to package N, along the longitudinal dimension, as shown in Figure 2. Each cell package includes two cells located on lane x and lane $\bar{x}$, denoted as cell (i, x) and cell $(i, \bar{x})$, respectively. It is assumed that there is only one bottleneck (caused by

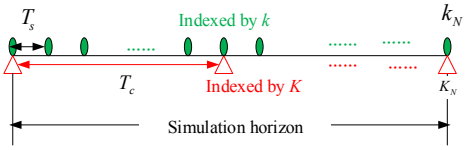


Figure 3: Relationship between different time horizons

on-ramp merging, lane drop or traffic accident) on this freeway segment, which is located at the end of cell (N, x). An abstract formulation of the optimal control problem can be thus defined.

$$\min J_p\left(\mathbf{x}(k_N)\right) + \sum_{k=1}^{k_N} J\left(\mathbf{x}(k), \mathbf{q}(k)\right) \tag{1}$$

subject to traffic flow conservation

$$\mathbf{x}(k+1) = \text{fun}\left(\mathbf{x}(k), \mathbf{q}(k), \mathbf{c}(K)\right), \quad 1 < k < k_N,$$

definition constraints to control variables

$$\mathbf{0} \le \mathbf{c}(K) \le \mathbf{C}, \text{ for } 1 < K < \lceil k_N \cdot T_s / T_c \rceil$$

and physical definition constraints to state variables and flow variables

$$\begin{cases} \mathbf{A}_1 \cdot \mathbf{q}(k) = \mathbf{B}_1 \\ \mathbf{A}_2 \cdot \mathbf{q}(k) \le \mathbf{B}_2 \\ \mathbf{0} \le \mathbf{x}(k) \le \mathbf{X}(k) \\ \mathbf{0} \le \mathbf{q}(k) \le \mathbf{Q}(k) \end{cases}, \quad 1 < k < k_N$$

where the objective function will be specified in Equation (2) with $k = 1, \cdots, k_N$ as the simulation steps and $K = \lceil k \cdot T_s / T_c \rceil$. T_s denotes the time duration of a simulation step. $\lceil g \rceil$ is the ceiling function that outputs the least integer greater than or equal to g. Figure 3 demonstrates the relationship between the simulation step, control cycle, and simulation horizon. To simplify the notation, we have used CA (as CAV) and RH (as RHV) for vehicle classes, while D (as DLC)[1] and M (as MLC)[2] in the equations throughout the paper. The notations involved in defining the optimal control problem, and the constraints are defined and summarized in Table 1.

The term fun() is a compact representation of the underlying traffic flow model that describes the multiclass multilane traffic dynamics. $\mathbf{x}(k)$ is the vector of state variables including $\rho_{i,x,d}(k)$, $v_{i,x}(k)$, and $W_{x,d}(k)$, during $[kT_s, (k+1)T_s)$. $\mathbf{q}(k)$ is the vector of sending and receiving functions that govern flow propagation in the multiclass multilane traffic flow model during $[kT_s, (k+1)T_s)$. The propagation of $\mathbf{x}(k)$ and the evaluation of $\mathbf{q}(k)$ will be described in the forthcoming section.

$\mathbf{c}(K)$ is the vector of control variables (which are the decision variables). If the optimal control problem is solvable, the optimal control strategy $\mathbf{c}^*(K)$ will be implemented during the time period $[KT_c, (K+1)T_c)$, where T_c denotes the duration of each control cycle and the integer $\lceil T_c / T_s \rceil$ is the number of simulation steps within each control cycle.

$$\begin{aligned} J\left(\mathbf{x}(k), \mathbf{q}(k)\right) = & \underbrace{\sum_{i=1}^{N} \sum_{y=\{x,\,\bar{x}\}} \sum_{d=\{CA,\,RH\}} V_t \cdot \left(\rho_{i,\,y,\,d}(k) \cdot l_i + W_{y,\,d}(k) + W_{on,\,d}(k)\right) \cdot T_s}_{\textit{Total time spent+Waiting time of extra-queue and on the ramps}} \\ & + \underbrace{\sum_{i=1}^{N} \sum_{y=\{x,\,\bar{x}\}} \sum_{Z=\{1,2,3\}} c_f \cdot F_{Z,\,i,\,y,\,RH}(k) \cdot e^{\left(V_{Z,\,i,\,y}(k) \cdot \mathrm{K}_f \cdot A_{Z,\,i,\,y}(k)^T\right)} \cdot T_s}_{\textit{Fuel consumption of RHV}} \\ & + \underbrace{\sum_{i=1}^{N} \sum_{e=CO,HC,NO} \sum_{y=\{x,\,\bar{x}\}} \sum_{Z=\{1,2,3\}} c_e \cdot F_{Z,\,i,\,y,\,RH}(k) \cdot e^{\left(V_{Z,\,i,\,y}(k) \cdot \mathrm{K}_e \cdot A_{Z,\,i,\,y}(k)^T\right)} \cdot T_s}_{\textit{Emission cost of RHV}} \\ & + \underbrace{\sum_{i=1}^{N} \sum_{y=\{x,\,\bar{x}\}} \sum_{Z=\{1,2,3\}} c_c \cdot F_{Z,\,i,\,y,\,CA}(k) \cdot e^{\left(V_{Z,\,i,\,y}(k) \cdot \mathrm{K}_c \cdot A_{Z,\,i,\,y}(k)^T\right)} \cdot T_s}_{\textit{Electricity cost of CAV}} \end{aligned} \tag{2}$$

$$J_p\left(\mathbf{x}(k_N)\right) = \varphi \cdot \sum_{y=\{x,\,\bar{x}\}} \underbrace{V_t \cdot \left(W_{y,\,CA}(k_N) + W_{y,\,RH}(k_N)\right)}_{\textit{Penalty on the residual extra-queue}} + \varphi \cdot \sum_{y=\{x,\,\bar{x}\}} \underbrace{V_t \cdot \left(W_{on,\,CA}(k_N) + W_{on,\,RH}(k_N)\right)}_{\textit{Penalty on the residual on-ramps}}$$

[1] Discretionary lane changing (DLC) is executed when the driver seeks for better driving conditions to gain a speed or travel time advantage.
[2] Mandatory lane changing (MLC) is executed when a driver must change lane to follow a certain path to his/her destination.

Table 1: Key variables of the optimal control problem

Element	Notation description	Duration	Constraint
$\rho_{i,x,\mathrm{CA}}(k)$	Density of CAVs on cell $(i,\ x)$ during step k		
$\rho_{i,x,\mathrm{RH}}(k)$	Density of RHVs on cell $(i,\ x)$ during step k		$\left[0,\ \frac{1}{l+C}\right]$
$P_{i,x,\mathrm{CA}}(k)$	Proportion/Penetration rate of CAVs on cell $(i,\ x)$ during step k		$[0,\ 1]$
$v_{i,x}(k)$	Traffic speed on cell $(i,\ x)$ during step k		$[0,\ \hat{v}_{i,x}(K)]$
$W_{\mathrm{on},i,\mathrm{CA}}\ (k)$	Queuing delay at the on-ramp connecting to cell i		
$W_{\mathrm{on},i,\mathrm{RH}}\ (k)$	with respect to CAV or RHV during step k		
$s^{i+1,\bar{x}}_{i,x,\mathrm{CA}}(k)$	Sending function of CAVs issued by LCC from cell $(i,\ x)$ to cell $(i{+}1,\ \bar{x})$ during step k		
$s^{i+1,\bar{x}}_{M,i,x,\mathrm{RH}}(k)$	Sending function of RHVs with MLC or DLC	T_s	$[0, Q_{i,x}(k)\cdot P_{i,x,\mathrm{RH}}(k)]$
$s^{i+1,\bar{x}}_{D,i,x,\mathrm{RH}}(k)$	demand from cell $(i,\ x)$ to cell $(i{+}1,\ \bar{x})$ during step k		
$s_{st,i,x,\mathrm{CA}}(k)$	Sending function that intend to leave cell $(i,\ x)$ to cell		$[0, Q_{i,x}(k)P_{i,x,\mathrm{CA}}(k)]$
$s_{st,i,x,\mathrm{RH}}(k)$	$(i+1,\ x)$ during step k of CAV or RHV		$[0, Q_{i,x}(k)P_{i,x,\mathrm{RH}}(k)]$
$\tilde{H}^{i+1,\bar{x}}_{i,x,\mathrm{CA}}(k)$	Minimum space headway criteria for $s^{i+1,\bar{x}}_{i,x,\mathrm{CA}}(k)$		
$\tilde{H}^{i+1,\bar{x}}_{\mathrm{M},i,x,\mathrm{RH}}(k)$	Minimum space headway criteria for $s^{i+1,\bar{x}}_{\mathrm{M},i,x,\mathrm{RH}}(k)$ or		
$\tilde{H}^{i+1,\bar{x}}_{\mathrm{D},i,x,\mathrm{RH}}(k)$	$s^{i+1,\bar{x}}_{\mathrm{D},i,x,\mathrm{RH}}(k)$		
$q^{i+1,\bar{x}}_{i,x,\mathrm{CA}}(k)$	The amount of $s^{i+1,\bar{x}}_{i,x,\mathrm{CA}}(k)$ received by the target cell (or executed LCC)		$\left[0, s^{i+1,\bar{x}}_{i,x,\mathrm{CA}}(k)\right]$
$q^{i+1,\bar{x}}_{\mathrm{M},i,x,\mathrm{RH}}(k)$	The amount of $s^{i+1,\bar{x}}_{\mathrm{M},i,x,\mathrm{RH}}(k)$ and $s^{i+1,\bar{x}}_{\mathrm{D},i,x,\mathrm{RH}}(k)$ received by		$\left[0,\ s^{i+1,\bar{x}}_{M,i,x,\mathrm{RH}}(k)\right]$
$q^{i+1,\bar{x}}_{\mathrm{D},i,x,\mathrm{RH}}(k)$	the target cells, respectively		$\left[0, s^{i+1,\bar{x}}_{D,i,x,\mathrm{RH}}(k)\right]$
$R_{i,x}(k)$	Receiving function of cell $(i,\ x)$ during step k		$[0, Q_{i,x}(k)]$
$\hat{v}_{i,x}(K)$	Implemented speed limit on cell $(i,\ x)$ during control cycle K for VSLC and VSLR	T_c	v_f
$\hat{p}^{i+1,\bar{x}}_{i,x,\mathrm{CA}}(K)$	The flow of CAVs that are issued to switch from cell $(i,\ x)$ to cell $(i{+}1,\ \bar{x})$ during control cycle K for LCC		1
$\hat{B}^{\bar{x}}_{i,x}(K)$	Variable denoting whether the LC recommendation from cell $(i,\ x)$ to lane $\bar{x}$ is released during control cycle K for LCR		0 (not released) or 1 (released)
$\hat{L}(K)$	Duration of the red phase of control cycle K for RM		Integers in $[0, T_c]$

Table 2: Speed and acceleration in evaluating the fuel consumption, emission and electricity cost during $[kT_s, (k+1)\,T_s)$

Z	$F_{Z,i,x,d}(k)$	$v_{Z,i,x}\ (k)$	$a_{Z,i,x}\ (k)$
1	$q_{st,i,x,d}(k)\cdot T_s$	$v_{i,x}\ (k) \rightarrow v_{i+1,x}\ (k+1)$	$\frac{v_{i+1,x}(k+1)-v_{i,x}(k)}{T_s}$
2	$\left(q^{i+1,\bar{x}}_{M,i,x,d}(k) + q^{i+1,\bar{x}}_{D,i,x,d}(k)\right)T_s$	$v_{i,x}\ (k) \rightarrow v_{i+1,\bar{x}}\ (k+1)$	$\frac{v_{i+1,\bar{x}}(k+1)-v_{i,x}(k)}{T_s}$
3	$S_{\ i,x,d}(k)\cdot T_s - \sum_{Z=\{1,2\}} F_{Z,i,x,d}(k)$	$v_{i,x}\ (k) \rightarrow v_{i,x}\ (k+1)$	$\frac{v_{i,x}(k+1)-v_{i,x}(k)}{T_s}$

As specified in Equation (2), the objective function includes the total time spent, fuel consumption and emission cost of RHVs, electricity cost of CAVs, and the penalty on residual extra-queue within the planning horizon. The following parts of the objective function are converted into monetary costs:

1. The first linear term represents the total travel time (TTT) (converted to monetary cost) spent on the freeway stretch and the time spent in the time spent queuing at on-ramps. V_t is the value of time that converts the delay into monetary cost.

2. The second to the fourth terms represent the fuel consumption and emission cost associated with RHVs, and electricity cost of CAVs. These terms are closely related to speed, inter-cell and cross-cell accelera-

Table 3: Variables in traffic flow propagation

Vector	Element	Notation description
$\mathbf{f}(k)$	$\hat{v}_{i,x}(K)$	Speed limit control on cell (i,x) during cycle K
	$\rho^{cr}_{i,x}(k)$	Critical density on cell $(i,\ x)$ at step k
	$Q_{i,x}(k)$	Capacity on cell $(i,\ x)$ at step k
	$w_{i,x}(k)$	Wave-back speed on cell $(i,\ x)$ at step k
	$\rho^{J}_{i,x}(k)$	Jams density on cell $(i,\ x)$ at step k
$\mathbf{u}(k)$	$S_{0,x}(k)$	Inflow demand of lane x at step k
	$S_{0,on}(k)$	Inflow demand at on ramp step k
	$P_{0,x,\mathrm{CA}}(k)$	Proportion of CAVs associated with $S_{0,\ x}(k)$
	$P_{0,on,\mathrm{CA}}(k)$	Proportion of CAVs associated with $S_{0,on}(k)$
	$R_{E,x}(k)$	Available space of the downstream section at step k

tion/deceleration. Traffic flows associated with cell (i, x), are categorized into three items (Z=1, 2, 3) according to the moving directions, i.e., the straightforward flow $F_{1,\ i,\ x,\ d}(k)$ which enters cell $(i+1, x)$ from cell (i, x) during time interval $[kT_s, (k+1)\,T_s)$ with d stands for RHV and CAV, respectively; the lane-changing flow $F_{2,\ i,\ x,\ d}(k)$ that switches lane from cell (i, x) to cell $(i+1, \bar{x})$, where $\bar{x}$ denotes the lanes adjacent to lane x; and $F_{3,\ i,\ x,\ d}(k)$ the flow that retains its current cell-lane. c_f (USD/liter); c_e (USD/gram) and c_c (USD/kWh) are coefficients that convert the fuel, emission and electricity costs into the monetary counterpart. The vector of speed $V_{Z,\ i,\ x}(k) = \left[1,\ \ v_{Z,\ i,\ x}(k), v^2_{Z,\ i,\ x}(k),\ v^3_{Z,\ i,\ x}(k)\right]$ and the vector of acceleration/deceleration $A_{Z,\ i,\ x}(k) = \left[1,\ \ a_{Z,\ i,\ x}(k), a^2_{Z,\ i,\ x}(k),\ a^3_{Z,\ i,\ x}(k)\right]$ are necessary when calculating the fuel consumption rate (liter/s), emission rate (gram/s) and electricity consumption rate (watt/s). The elements of $V_{Z,\ i,\ x}(k)$ and $A_{Z,\ i,\ x}(k)$ are listed in Table 2 (Liu et al., 2017). $\mathrm{K}_f, \mathrm{K}_e$ and K_c are the model coefficient matrices (Ahn and Cassidy, 2007; Tang et al., 2017), to which examples will be depicted in Section 5.

3. The last term $J_p\,(\mathbf{x}(k_N))$ penalizes the extra-queue. The factor φ is set to be a large value, e.g., 1000, to eliminate residual queue at the end of the planing horizon.

Finally, the constraints $\mathbf{A}_1 \cdot \mathbf{q}(k)$=$\mathbf{B}_1,\ \mathbf{A}_2 \cdot \mathbf{q}(k) \leq \mathbf{B}_2$ will be depicted in the forthcoming section.

3. The multiclass multilane traffic flow model

3.1. Traffic-state propagation of flow mixed with CAVs and RHVs

The multiclass multilane cell transmission model (CTM) proposed by Pan et al. (2021) incorporates several new features. First, it considers the effect of time-dependent proportion of CAVs on the minimum safe gap acceptance and the cell-lane fundamental diagram, which are recapitulated in the appendix. Second, it defines the lane-level reaction laws of sending and receiving functions that govern the flow propagation with respect to the multiclass cell-lane fundamental diagram. And third, it enables lane-changing priorities with respect to different vehicle classes and the lane-changing motivations while considering the drivers' behavior. However, to adapt to the control design, we need to extend the model to consider the constraints induced by the control schemes and the heterogenous resolutions of simulation step k and control cycle K.

Take cell $(i,\ x)$ as shown in Figure 2 as an example, based on flow conservation, the cell densities $\rho_{i,x,\mathrm{CA}}(k+1)$ (P.C.U./mile/lane) of CAVs and $\rho_{i,x,\mathrm{RH}}(k+1)$ (P.C.U./mile/lane) of RHVs, are estimated as Equation (3).

$$\begin{cases} \rho_{i,\ x,\ \mathrm{CA}}(k+1) = \rho_{i,\ x,\ \mathrm{CA}}(k) \ + \frac{T_s}{l_i}\left(q^{i,\ x}_{st,\ i-1,\ x,\ \mathrm{CA}}(k) + q^{i,\ x}_{i-1,\ \bar{x},\mathrm{CA}}(k)\right) - \frac{T_s}{l_i}\left(q^{i+1,\ x}_{st,\ i,\ x,\ \mathrm{CA}}(k) + q^{i+1,\ \bar{x}}_{i,\ x,\ \mathrm{CA}}(k)\right) \\ \rho_{i,\ x,\ \mathrm{RH}}(k+1) = \rho_{i,\ x,\ \mathrm{RH}}(k) + \frac{T_s}{l_i}\left(q^{i,\ x}_{st,\ i-1,\ x,\ \mathrm{RH}}(k) + q^{i,\ x}_{M,\ i-1,\ \bar{x},\ \mathrm{RH}}(k) + q^{i,\ x}_{D,\ i-1,\ \bar{x},\ \mathrm{RH}}(k)\right) \\ \qquad\qquad\qquad\qquad - \frac{T_s}{l_i}\left(q^{i+1,\ x}_{st,\ i,\ x,\ \mathrm{RH}}(k) + q^{i+1,\bar{x}}_{M,\ i,\ x,\ \mathrm{RH}}(k) + q^{i+1,\ \bar{x}}_{D,\ i,\ x,\ \mathrm{RH}}(k)\right) \end{cases} \tag{3}$$

The estimation of density on cell $(i, \bar{x})$ can be similarly defined. The cell density is a summation of the two vehicle

classes, *i.e.*, $\rho_{i,\,x}(k+1)=\rho_{i,\,x,\,\mathrm{CA}}(k+1)+\rho_{i,\,x,\,\mathrm{RH}}(k+1)$. The proportion of CAVs in a particular cell-lane is given by

$$\begin{aligned} &P_{i,\,x,\,\mathrm{CA}}(k+1)=\rho_{i,\,x,\,\mathrm{CA}}(k+1)/\rho_{i,\,x}(k+1),\ \text{if}\ \ \rho_{i,\,x}(k+1)>0\\ &P_{i,\,x,\,\mathrm{RH}}(k+1)=1-P_{i,\,x,\,\mathrm{CA}}\,(k+1) \end{aligned} \tag{4}$$

while the cell-lane speed is given by the following equation.

$$v_{i,\,x}(k+1)=\begin{cases} \hat{v}_{i,\,x}(K)\ ,\ if\ \rho_{i,\,x}(k+1)\le\rho^{cr}_{i,\,x}(k) \\ \dfrac{\left(1-l\cdot\rho_{i,\,x}(k+1)-C\cdot\rho_{i,\,x}(k+1)\right)}{\rho_{i,\,x}(k+1)\left(P_{i,\,x,\,\mathrm{CA}}(k+1)\,\Delta T_{\mathrm{CA}}\ +P_{i,\,x,\,\mathrm{RH}}(k+1)\,\Delta T_{\mathrm{RH}}\right)}, \text{else} \end{cases} \tag{5}$$

depending on whether the speed limit control or the permanent compulsory speed limit is implemented (refer to the appendix for details). $\rho^{cr}_{i,\,x}(k)$ denotes the critical density of cell (i, x).

$$\rho^{cr}_{i,\,x}\,(k)=\frac{1}{\hat{v}_{i,\,x}\,(K)\left(P_{i,\,x,\,\mathrm{CA}}\,(k)\,\Delta T_{CA}+\left(1-P_{i,\,x,\,\mathrm{CA}}\,(k)\right)\Delta T_{RH}\right)+l+C} \tag{6}$$

ΔT_{CA} and ΔT_{RH} denote the response times of CAVs and RHVs, respectively. l (mile) denotes the vehicle length of the leading vehicle, and C is the safety gap or minimum safety constant gap when all the vehicles are at a standstill (Hidas, 2005; Jepsen, 1998). This lane-specific fundamental diagram was derived by Pan et al. (2021). To make the paper self-contained, we recapitulate the essence in the Appendix.

The evaluation of the vector of sending and receiving functions $\mathbf{q}(k)$ is vital for the propagating the multiclass multilane traffic flow for determining the state vector $\mathbf{x}(k)$. Apart from the state vector $\mathbf{x}(k)$ and the vector of the control variables $\mathbf{c}(K)$, the external input $\mathbf{u}(k)$ in conjunction with the fundamental diagram $\mathbf{f}(k)$ is also necessary for the evaluation of $\mathbf{q}(k)$. To this end, we represent $\mathbf{q}(k)$ as a mapping of these vectors as follows:

$$\mathbf{q}(k)=\mathrm{fun}_q(\mathbf{x}(k),\mathbf{c}(K),\mathbf{u}(k),\mathbf{f}(k)),\ \text{for}\ 1<k<k_N,\ K=\lceil k\cdot T_s/T_c\rceil$$

which can be regarded as equality constraints of the optimization problem.

3.1.1. Sending flow and lane changing demand functions

The sending function $s^{i+1}_{i,\,x,\,d}\,(k)$ (P.C.U./hour/lane), which quantifies the traffic flow that intends to leave cell $(i,\,x)$ for vehicle class d, is defined as

$$s^{i+1}_{i,\,x,\,d}\,(k)=\begin{cases} \hat{v}_{i,\,x}(K)\cdot\rho_{i,\,x}(k)\cdot P_{i,\,x,\,d}(k), & \text{if}\ \ \rho_{i,\,x}(k)<\rho_{i,\,x,\,c}(k) \\ Q_{i,\,x}(k)\cdot P_{i,\,x,\,d}(k), & \text{if}\ \rho_{i,\,x}(k)\ge\rho_{i,\,x,\,c}(k) \end{cases} \tag{7}$$

for $i=1,2,\cdots N-1$, *and on*; equals to $S_{0,\,x}(k)\cdot P_{0,\,x,\,d}(k)+W_{x,\,d}(k)$, if $i=0$; and equals to $S_{0,\,on}(k)\cdot P_{0,\,on,\,d}(k)+W_{on,\,d}(k)$, if it connects to the on-ramp, i.e., $i=0$, *on*. $P_{i,\,x,\,d}(k)$ denotes the proportions of CAVs and RHVs in cell $(i,\,x)$. $P_{0,\,x,\,d}(k)$ is the proportion of vehicle class d arriving at the upstream boundary, $S_{0,\,x}(k)$ is the demand toward lane x and $W_{x,\,d}(k)$ is the number of vehicles of class d queuing at the upstream boundary to cell $(1,\,x)$ at time step k.

The sending function $s^{i+1}_{i,\,x,\,d}\,(k)$ determines the flow intended to leave cell (i, x) during time step k and move toward the downstream cell package $(i+1)$ involving all possible lanes, i.e., cell $(i{+}1,\bar{x})$ in the adjacent lane and cell $(i{+}1,\,x)$ in the current lane. The lane changing demands can be further categorized into three possible types: a) CAVs following the lane-changing instructions issued by VACS, which is denoted by $s^{i+1,\,\bar{x}}_{i,\,x,\,\mathrm{CA}}(k)$; b) RHVs making MLC decisions based on lane-changing suggestions, which is denoted by $s^{i+1,\,\bar{x}}_{M,\,i,\,x,\,\mathrm{RH}}(k)$; and c) RHVs making DLC decisions to gain speed advantage, which is denoted by $s^{i+1,\,\bar{x}}_{D,\,i,\,x,\,\mathrm{RH}}(k)$. These lane-changing flows can be evaluated as

$$\begin{aligned} &s^{i+1,\bar{x}}_{i,\,x,\,\mathrm{CA}}(k)=s^{i+1}_{i,\,x,\,\mathrm{CA}}(k)\cdot\hat{p}^{i+1,\,\bar{x}}_{i,\,x,\,\mathrm{CA}}(K)\\ &s^{i+1,\,\bar{x}}_{M,\,i,\,x,\,\mathrm{RH}}(k)=s^{i+1}_{i,\,x,\,\mathrm{RH}}(k)\cdot p^{i+1,\,\bar{x}}_{M,\,i,\,x,\,\mathrm{RH}}(k)\cdot\hat{B}^{\bar{x}}_{i,\,x}(K)\\ &s^{i+1,\,\bar{x}}_{D,\,i,\,x,\,\mathrm{RH}}(k)=s^{i+1}_{i,\,x,\,\mathrm{RH}}(k)\cdot p^{i+1,\,\bar{x}}_{D,\,i,\,x,\,\mathrm{RH}}(k)\cdot\left(1-\hat{B}^{\bar{x}}_{i,\,x}(K)\right) \end{aligned} \tag{8}$$

$p^{i+1,\bar{x}}_{M,\,i,\,x,\,\mathrm{RH}}(k)$ and $p^{i+1,\,\bar{x}}_{D,\,i,\,x,\,\mathrm{RH}}(k)$ are associated with human drivers' decisions in response to the LCR. The term $p^{i+1,\bar{x}}_{M,\,i,\,x,\,\mathrm{RH}}(k)$

evaluated by Equation (9) depends on the following elements: the remaining distance[3] d_i from the current cell package i to the target turning point; the cell density on target lane $\rho_{i,\bar{x}}(k)$; the cumulative MLC demand that intends to switch from lane x to lane $\bar{x}$ along the stretch, i.e., $\hat{B}^{\bar{x}}_{i,x}(K)\cdot S^{\bar{x}}_{M,\,x,\,\mathrm{RH}}(k)$; and the accumulation of executed MLC flow from cell $(1,\,x)$ to cell $(i,\,x)$. Interested readers are referred to Pan et al. (2016, 2021) for a detailed discussion on this and the calibration of parameters α_1, α_2 and the critical remaining distance d_c. **This term describes the drivers' compliance with the lane-changing recommendation (LCR)**.

$$p^{i+1,\,\bar{x}}_{M,\,i,\,x,\,\mathrm{RH}}(k)=\left[\hat{B}^{\bar{x}}_{i,x}(K)\cdot s^{1,\,\bar{x}}_{M,\,0,\,x,\,\mathrm{RH}}(k)\cdot \mathrm{e}^{-(d_i-d_c)^2/(\alpha_1+\alpha_2\cdot\rho_{i,\bar{x}}(k))^2}-\sum_{j=1}^{i-1} q^{j+1,\bar{x}}_{M,\,j,\,x,\mathrm{RH}}(k)\right]/s^{i+1}_{i,\,x,\,\mathrm{RH}}(k) \tag{9}$$

The DLC is motivated by the speed advantage of the adjacent lane if no LCR is issued to current segment on current lane, i.e., $\hat{B}^{\bar{x}}_{i,\,x}(K)=0$.

$$p^{i+1,\,\bar{x}}_{D,\,i,\,x,\,\mathrm{RH}}(k)=\frac{\max\left(0,\ v_{i,\,\bar{x}}(k)-v_{i,\,x}(k)\right)}{v_{f,\,i}\cdot\tau} \tag{10}$$

The term τ can be interpreted as the average time a driver takes to decide and execute a lane change when the original lane is stopped and the target lane is free-flowing (Pan et al., 2016).

Finally, the straightforward flow intended to leave cell $(i,\,x)$ and enter cell $(i{+}1,\,x)$ is

$$\begin{aligned} s^{i+1,\,x}_{st,\,i,\,x,\mathrm{CA}}(k)&=s^{i+1}_{i,\,x,\,\mathrm{CA}}(k)-s^{i+1,\,\bar{x}}_{i,\,x,\,\mathrm{CA}}(k)\\ s^{i+1,x}_{st,\,i,\,x,\,\mathrm{RH}}(k)&=s^{i+1}_{i,\,x,\,\mathrm{RH}}(k)-s^{i+1,\,\bar{x}}_{M,\,i,\,x,\,\mathrm{RH}}(k)-s^{i+1,\,\bar{x}}_{D,\,i,\,x,\,\mathrm{RH}}(k)\end{aligned} \tag{11}$$

For the on-ramp traffic, all vehicles have to merge to the mainstream cell $(N,\ x)$ from the ramp cell (on), i.e., executing mandatory lane-changing. Considering the ramp traffic signal, the sending function from the on-ramp $s^{N,\,x}_{M,\,on}(k)$, (P.C.U./hour/lane) can be evaluated as follow:

$$s^{N,\,x}_{M,\,on}(k)=\begin{cases} v_{f,\,on}(k)\cdot\rho_{on}(k)\,(1-r_{rm}(k)) & if\ \rho_{on}(k)<\rho^{cr}_{on}(k)\\ w_{on}(k)\cdot\left(\rho^{J}_{on}(k)-\rho_{on}(k)\right)\cdot(1-r_{rm}(k)), & otherwise\end{cases} \tag{12}$$

where $r_{rm}(k)$ is thus defined for the red phase of the ramp signal:

$$r_{rm}(k)=\begin{cases} 1 & if\quad k-\left\lfloor\frac{k}{K}\right\rfloor\cdot K\le\hat{L}(K)\\ 0, & if\quad k-\left\lfloor\frac{k}{K}\right\rfloor\cdot K>\hat{L}(K)\end{cases} \tag{13}$$

with $\lfloor g\rfloor$ the floor function that outputs the largest integer less than or equal to g.

3.1.2. Receiving flow functions

However, the execution of all lane changing maneuvers and straightforward demands still depend on two factors: the fulfillment of minimum gap acceptance criteria and different merging priorities by the target lane. The receiving function of the target cell $(i{+}1,\bar{x})$ (P.C.U./hour/lane) can be evaluated according to the fundamental diagram $\mathbf{f}(k)$.

$$R_{i+1,\,\bar{x}}(k)=\begin{cases} Q_{i+1,\,\bar{x}}(k),\ \text{if}\ \rho_{i+1,\,\bar{x}}(k)<\rho^{cr}_{i+1,\,\bar{x}}(k)\\ w_{i+1,\,\bar{x}}(k)\left(\rho^{J}_{i+1,\,\bar{x}}(k)-\rho_{i+1,\,\bar{x}}(k)\right),\ \text{otherwise}\end{cases} \tag{14}$$

for $i=0,2,\cdots N-1$, and equals to $R_{E,\,\bar{x}}(k)$, for $i=N$ where E is the downstream sink connected to the last cell N.

As it can be inferred from Section 3.1, towards the target cell $(i{+}1,\bar{x})$, there are five sending flows, i.e., $s^{i+1,\,x}_{st,\,i,\,x,\,\mathrm{CA}}(k)$, $s^{i+1,\,x}_{st,\,i,\,x,\,\mathrm{RH}}(k)$, $s^{i+1,\,\bar{x}}_{i,\,x,\,\mathrm{CA}}(k)$, $s^{i+1,\,\bar{x}}_{M,\,i,\,x,\,\mathrm{RH}}(k)$, and $s^{i+1,\,\bar{x}}_{D,\,i,\,x,\,\mathrm{RH}}(k)$, where the last three items can be further categorized as

[3]The remaining distance d_i is defined as the distance from the current position of the subject vehicle to its target turning point. d_c and d_r denote the remaining distances by which the test section is partitioned as remote, median, and close sections, respectively, with each section corresponding to a specific level of mandatory lane-changing (MLC) urgency and minimum acceptance criterion. Taking a vehicle intending to execute an MLC as an example, the target turning point is considered to be remote as long as the remaining distance $d_i>d_r$ and close if $d_i<d_c$.

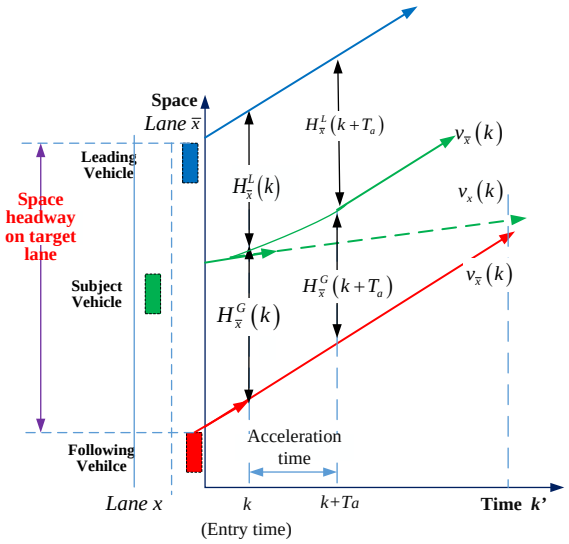


Figure 4: The headway acceptance analysis

$s_{i,\,x,\,\mathrm{CA}}^{i+1,\,\bar{x},\,f}(k)$, $s_{M,\,i,\,x,\,\mathrm{RH}}^{i+1,\,\bar{x},\,f}(k)$ and $s_{D,\,i,\,x,\,\mathrm{RH}}^{i+1,\,\bar{x},\,f}(k)$, according to the driving mode of following vehicle on target lane, and the superscript f denotes this item. The feasibility of these lane changing demands as mentioned above in merging to target lane are assessed via comparing the actual adjacent space headway $H_{i+1,\,\bar{x}}(k)$ with three minimum space headway criteria $\tilde{H}_{i,\,x,\,\mathrm{CA}}^{i+1,\,\bar{x},\,f}(k)$, $\tilde{H}_{M,\,i,\,x,\,\mathrm{RH}}^{i+1,\,\bar{x},\,f}(k)$, and $\tilde{H}_{D,\,i,\,x,\,\mathrm{RH}}^{i+1,\,\bar{x},\,f}(k)$, respectively.

3.1.3. The minimum space headway criteria

Apart from the traffic flow characteristics associated with traffic flow on a single lane, such as in the fundamental diagram and space headway distribution previously introduced, the penetration of CAVs also has an impact on lane-changing maneuvers when traveling on multilane freeways. Suppose that the subject vehicle traveling on lane x intends to switch to lane $\bar{x}$ via the space gap between the leading vehicle and the following vehicle in the target lane at time k as demonstrated in Figure 4. It is assumed that the subject vehicle is traveling at speed $v_x(k)$ (miles/h), whereas the leading vehicle and following vehicle in the target lane are both traveling at the same speed $v_{\bar{x}}(k)$(miles/h). Without loss of generality, the assumption $v_x(k) < v_{\bar{x}}(k)$ is proposed in this section, i.e., the lane-changing maneuver is executed to gain better driving conditions. Based on the speed difference assumption $v_x(k) < v_{\bar{x}}(k)$, it is necessary for the subject vehicle to accelerate to avoid a collision with the following vehicle in the target lane after the lane changing maneuver is executed. During the acceleration process, the space gap between the leading vehicle and the following vehicle must always fulfill the collision avoidance criteria.

Denoting $k' = k + T_a$ $(hour)$ as the moment when the lane-changing maneuver is executed, as shown in Figure 4, the subject vehicle is supposed to travel at the same speed as the vehicles in the target lane after acceleration. By assumption, this speed is $v_x(k+T_a) = v_{\bar{x}}(k+T_a) = v_{\bar{x}}(x)$. Considering the speed difference between the related vehicles during the lane-changing maneuver, the leading space headway[4] $H_{\bar{x}}^{L}(k')$ (mile) will gradually increase, whereas the lag space headway $H_{\bar{x}}^{G}(k')$ (mile) will decrease with respect to the acceleration and lane-changing maneuver of the subject vehicle, that is,

$$\begin{cases} H_{\bar{x}}^{L}(k') = H_{\bar{x}}^{L}(k) + (v_{\bar{x}}(k) - v_x(k))\,T_a - \frac{1}{2}aT_a^2 \\ H_{\bar{x}}^{G}(k') = H_{\bar{x}}^{G}(k) - (v_{\bar{x}}(k) - v_x(k))\,T_a + \frac{1}{2}aT_a^2 \end{cases} \tag{15}$$

and

$$\begin{cases} H_{\bar{x}}^{L}(k') = H_{\bar{x}}^{L}(k+T_a) \\ H_{\bar{x}}^{G}(k') = H_{\bar{x}}^{G}(k+T_a) \end{cases}$$

[4]The leading space headway denotes the distance from the head of the lead vehicle to the head of the subject vehicle. The lag headway denotes the distance from the head of the subject vehicle to the head of the following vehicle.

where a (miles/hour^2) is the average acceleration of the subject vehicle. According to the collision avoidance principle previously described, the following conditions need to be fulfilled:

$$\begin{cases} H_{\bar{x}}^{L}(k') \geq (v_x(k) + aT_a)\,\Delta T_s + l + C \\ H_{\bar{x}}^{G}(k') \geq v_{\bar{x}}(k)\,\Delta T_f + l + C \end{cases} \tag{16}$$

where ΔT_d and ΔT_f denote the response times of the subject and the following vehicles, respectively. The space headway of the target lane at time k, $H_{x,\,d}^{\bar{x},\,f}(k) = H_{\bar{x}}^{L}(k) + H_{\bar{x}}^{G}(k)$, should be not less than the minimum space headway $\tilde{H}_{lc,\,x,\,d}^{\bar{x},\,f}(k)$, which is determined by Equation (17).

$$H_{x,\,d}^{\bar{x},\,f}(k) \geq \tilde{H}_{lc,\,x,\,d}^{\bar{x},\,f}(k) = \begin{cases} \left[\underbrace{v_{\bar{x}}(k)\,\Delta T_f}_{item\ 1} + \underbrace{v_x(k)\,\Delta T_d}_{item\ 2} + \underbrace{\dfrac{(v_{\bar{x}}(k) - v_x(k))^2}{2a}}_{item\ 3} \right] + \underbrace{(2l + 2C)}_{item\ 4} & Case1 \\ \left[\underbrace{v_{\bar{x}}(k)\,\Delta T_f}_{item\ 1} + \underbrace{v_x(k)\,\Delta T_d}_{item\ 2} + \underbrace{\dfrac{(v_{\bar{x}}(k) - v_x(k))^2}{2a}}_{item\ 3} \right] \frac{d_i - d_c}{d_r - d_c} + \underbrace{(2l + 2C)}_{item\ 4} & Case2 \\ \underbrace{(2l + 2C)}_{item\ 4} & Case3 \end{cases} \tag{17}$$

As shown in Equation (17), the minimum space headway criteria are divided into three different cases with each applicable to lane changing scenarios regarding the lane-changing motivation, level of lane-changing urgency, and whether the gap control coordinates the lane changing (for CAVs):

Case 1 is applicable to the non-coordinated DLC or a non-coordinated non-urgent MLC which is proposed at a remote state with the remaining distance[5] $d_i > d_r$, and the following vehicle would not cooperate with this lane changing demand, i.e., the following vehicle is RHV, or the GC order is not issued to the following CAV. Under this case, drivers usually prefer a relatively large gap at the beginning of a lane-changing maneuver due to a risk-adverse attitude.

Case 2 refers to the non-coordinated MLC between non-urgent scenario (that the lane change is executed far from the target point due to the risk-adverse attitude of the drivers) and the urgent scenario (that the vehicle is so close to the target point that the lane change has to be executed), saying $d_c < d_i < d_r$. Under this case, $\tilde{H}_{lc,\,x,\,d}^{\bar{x},\,f}(k)$ decreases linearly with respect to the remaining distance when d is within the range $d_c \leq d_i \leq d_r$ according to Case 2 of Equation (17). Both Case 1 and Case 2 were devised to model the minimum gap acceptance criterion of RHVs in line with Pan et al. (2016).

Case 3 involves the coordinated DLC and MLC of the CAVs, or the non-coordinated urgent MLC, i.e., $d_i < d_c$.

Whether the target lane could provide sufficient space headway to fulfill the minimum space headway criteria is a prerequisite of a successful lane-changing maneuver. For traffic flow of RHVs, Zheng et al. (2013) quantified the impact of lane-changing maneuvers on the immediately following vehicle using Newell's car following model. Compared with traffic of pure RHVs, traffic flow mixed with CAVs would introduce the following three advantages: 1) the higher penetration rate of CAVs induces a higher capacity for the same freeway; 2) the minimum space headway gap required by the CAVs is smaller than that required by RHVs, for both MLC and DLC lane changing intentions; 3) the space headway in front of RHVs can be significantly enhanced if more CAVs are traveling on the segment under the same traffic conditions because CAVs tolerate a smaller headway. As it can be inferred from advantages 2) and 3), lane changing will be easier and safer. The V2V system plays an important role in measuring the (relative) speed, acceleration, and gaps between adjacent vehicles and lanes to support the minimum safety gap control and the lane changing control process.

[5]The remaining distance d_i is defined as the distance from the current position of the subject vehicle to its target turning point. d_c and d_r denote the remaining distances by which the test section is partitioned as remote, median, and close sections, respectively, with each section corresponding to a specific level of mandatory lane-changing (MLC) urgency and minimum acceptance criterion. Taking a vehicle intending to execute an MLC as an example, the target turning point is considered to be remote as long as the remaining distance $d_i > d_r$ and close if $d_i < d_c$.

3.1.4. Execution of lane changing maneuvers

Considering the limited available space provided by the target cell, the flows that can be actually received by the target cell are calculated by Equations (18) and (19). The subscript lc in $q^{i+1,\,\bar{x}}_{lc,\,i,\,x,\,d}(k)$ (P.C.U./hour/lane) refers to the lane-changing motivations and d refers to vehicle types of subject vehicle. $U^{i+1,\,\bar{x}}_{i}(k)$ (P.C.U./hour/lane) denotes the total space required by the lane-changing demand toward cell $(i{+}1,\bar{x})$ fulfilling the minimum space headway criterion and extra spaces in the target lane.

$$q^{i+1,\,\bar{x},\,f}_{lc,\,i,\,x,\,d}(k) = \begin{cases} s^{i+1,\,\bar{x},\,f}_{lc,\,i,\,x,\,d}(k) & \text{if } U^{i+1,\,\bar{x}}_{i}(k) \le R_{i+1,\,\bar{x}}(k) \text{ and } \tilde{H}^{i+1,\,\bar{x},\,f}_{lc,\,i,\,x,\,d}(k) \le H_{i+1,\,\bar{x},\,f}(k) \\ \dfrac{s^{i+1,\,\bar{x},\,f}_{lc,\,i,\,x,\,d}(k)}{U^{i+1,\,\bar{x}}_{i}(k)} R_{i+1,\bar{x}}(k) & \text{if } U^{i+1,\,\bar{x}}_{i}(k) > R_{i+1,\,\bar{x}}(k) \text{ and } \tilde{H}^{i+1,\,\bar{x},\,f}_{lc,\,i,\,x,\,d}(k) \le H_{i+1,\,\bar{x},\,f}(k) \\ 0 & \text{if } \tilde{H}^{i+1,\,\bar{x},\,f}_{lc,\,i,\,x,\,d}(k) > H_{i+1,\,\bar{x},\,f}(k) \end{cases} \tag{18}$$

$$q^{i+1,\,\bar{x}}_{st,\,i,\,\bar{x},\,d}(k) = \begin{cases} s^{i+1,\,\bar{x}}_{st,\,i,\,\bar{x},\,d}(k) & \text{if } U^{i+1,\,\bar{x}}_{i}(k) \le R_{i+1,\,\bar{x}}(k) \\ \dfrac{s^{i+1,\,\bar{x}}_{st,\,i,\,\bar{x},d}(k)}{U^{i+1,\bar{x}}_{i}(k)} R_{i+1,\,\bar{x}}(k) & \text{if } U^{i+1,\,\bar{x}}_{i}(k) > R_{i+1,\,\bar{x}}(k) \end{cases} \tag{19}$$

The actual space headway target cell$(i{+}1,\ \bar{x})$ in front of vehicle with $f = CAV$ or $f = RHV$ is calculated as :

$$\begin{aligned} H_{i+1,\,\bar{x},\,\mathrm{CA}}(k) &= \begin{cases} v_{\,i+1,\,\bar{x}}(k)\,\Delta T_{\mathrm{CA}} + l + C, & \text{if } \rho_{i+1,\bar{x}}(k) \ge \rho^{cr}(k) \\ \dfrac{v_{\,i+1,\,\bar{x}}(k)\Delta T_{\mathrm{CA}} + l + C}{\rho_{i+1,\bar{x}}(k)\cdot(l+C) + \rho_{i+1,\bar{x}}(k)\cdot v_{\,i+1,\,\bar{x}}(k)\cdot\left(P_{i+1,\bar{x},\,CA}(k)\Delta T_{\mathrm{CA}} + \left(1 - P_{i+1,\bar{x},\,CA}(k)\right)\Delta T_{\mathrm{RH}}\right)}, & \text{otherwise} \end{cases} \\ H_{i+1,\,\bar{x},\,\mathrm{RH}}(k) &= \begin{cases} v_{\,i+1,\,\bar{x}}(k)\,\Delta T_{\mathrm{RH}} + l + C & \text{if } \rho_{i+1,\,\bar{x}}(k) \ge \rho^{cr}_{i+1,\,\bar{x}}(k) \\ \dfrac{v_{\,i+1,\,\bar{x}}(k)\Delta T_{\mathrm{RH}} + l + C}{\rho_{i+1,\,\bar{x}}(k)\cdot(l+C) + \rho_{i+1,\,\bar{x}}(k)\cdot v_{\,i+1,\,\bar{x}}(k)\cdot\left(P_{i+1,\bar{x},\,CA}(k)\Delta T_{\mathrm{CA}} + \left(1 - P_{i+1,\bar{x},\,CA}(k)\right)\Delta T_{\mathrm{RH}}\right)}, & \text{otherwise} \end{cases} \end{aligned} \tag{20}$$

And the total space expected by all the merging flows is evaluated as $U^{i+1,\,\bar{x}}_{i}(k)$:

$$U^{i+1,\,\bar{x}}_{i}(k) = s^{i+1,\bar{x}}_{st,\,i,\,\bar{x},\,\mathrm{CA}}(k) + s^{i+1,\,\bar{x}}_{st,\,i,\,\bar{x},\,\mathrm{RH}}(k) + \sum_{lc=\{M,D\}} \sum_{d=\{\mathrm{CA},\,\mathrm{RH}\}} \sum_{f=\{\mathrm{CA},\,\mathrm{RH}\}} s^{i+1,\,\bar{x},\,f}_{lc,\,i,\,x,\,d}(k)\,\frac{\tilde{O}^{i+1,\,\bar{x},\,f}_{lc,\,i,\,x,\,d}(k)}{l},\ \forall \tilde{H}^{i+1,\,\bar{x},\,f}_{lc,\,i,\,x,\,d}(k) < H_{i+1,\,\bar{x},\,f}(k) \tag{21}$$

the minimum occupied space for executing the lane changing maneuver $\tilde{O}^{i+1,\,\bar{x},\,f}_{lc,\,i,\,x,\,d}(k)$ is evaluated as:

$$\tilde{O}^{i+1,\,\bar{x},\,f}_{lc,\,i,\,x,\,d}(k) = v_{i+1,\bar{x}}(k)\,\Delta T_f + v_{i,x}(k)\,\Delta T_d + \frac{\left(v_{i+1,\bar{x}}(k) - v_{i,x}(k)\right)^2}{2a} + (2l + 2C) \tag{22}$$

Assuming the driving mode of following vehicle follows the penetration rate of cell $(i+1,\ \bar{x})$, the final lane changing flows is estimated as follow:

$$\begin{aligned} q^{i+1,\,\bar{x}}_{i,\,x,\,CA}(k) &= q^{i+1,\,\bar{x},\,CA}_{i,\,x,\,CA}(k) P_{i+1,\,\bar{x},\,CA}(k) \ + \ q^{i+1,\,\bar{x},\,RH}_{i,\,x,\,CA}(k) P_{i+1,\,\bar{x},\,RH}(k) \\ q^{i+1,\,\bar{x}}_{D,\,i,\,x,\,RH}(k) &= q^{i+1,\,\bar{x},\,CA}_{D,\,i,\,x,\,RH}(k) P_{i+1,\,\bar{x},\,CA}(k) \ + q^{i+1,\,\bar{x},\,RH}_{D,\,i,\,x,\,RH}(k) P_{i+1,\,\bar{x},\,RH}(k) \\ q^{i+1,\,\bar{x}}_{M,\,i,\,x,\,RH}(k) &= \ q^{i+1,\,\bar{x},\,CA}_{M,\,i,\,x,\,RH}(k) P_{i+1,\,\bar{x},\,CA}(k) \ + q^{i+1,\,\bar{x},\,RH}_{M,\,i,\,x,\,RH}(k) P_{i+1,\,\bar{x},\,RH}(k) \end{aligned} \tag{23}$$

Finally, the number of vehicles queuing upstream at the boundary cell $(1,\ x)$ are evaluated as follows:

$$\begin{aligned} W_{x,\,\mathrm{CA}}(k+1) &= W_{x,\,\mathrm{CA}}(k) + \left(S_{0,\,x}(k) P_{0,\,x,\,\mathrm{CA}}(k) - q^{1,\,x}_{st,\,0,\,x,\,\mathrm{CA}}(k)\right) T_s \\ W_{x,\,\mathrm{RH}}(k+1) &= W_{x,\,\mathrm{RH}}(k) + \left(S_{0,\,x}(k) P_{0,\,x,\,\mathrm{RH}}(k) - q^{1,\,x}_{st,\,0,\,x,\,\mathrm{RH}}(k)\right) T_s \end{aligned}$$

3.2. Constraints on the control variables

As mentioned, we need to extend the multiclass multilane model proposed in Pan et al. (2021) to consider the constraints induced by the control schemes and the heterogenous resolutions of simulation step k and control cycle K to support the control design. A set of constraints on the control variables are listed in Equation (24). Considering the feasibility of the speed limit and to maintain traffic flow stability, $\hat{v}_{i,\,x}(K)$ is constrained on the lateral (the 1^{st} equation), longitudinal (the 2^{nd} equation and the 3^{rd} inequality), and temporal dimensions (the 4^{th} inequality), and it

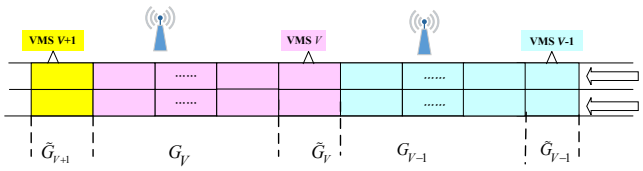


Figure 5: Illustration of freeway traffic control by VMS gantries

is discretized into several intervals with sufficient increments (rather than changing continuously, the 5^{th} constraint). Besides, $\hat{B}^{\bar{x}}_{i,\,x}(K)$, denoting whether the RHVs travelling on cell (i, x) is recommended to switch to the adjacent lane $\bar{x}$, is also constrained on the lateral (the 6^{th} equation), longitudinal (the 7^{th} equation), and their changes are restricted to be less than 2 times (the 8^{th} inequality). The LCC for CAVs is similarly restricted (the 10^{th} equation). The LC ratio $\hat{p}^{i+1,\,\bar{x}}_{i,\,x,\,\mathrm{CA}}(K)$ of each cell should be within 0%-100% (the 9^{th} constraint). Finally, the duration time $\hat{L}(K)$ of red light for ramp metering is restricted as an integer between 0 and control cycle time T_c (the 11^{th} and 12^{th} constraints).

$$
\begin{aligned}
&1)\ \left|\hat{v}_{i,\,x}(K)-\hat{v}_{i,\,\bar{x}}(K)\right| \le \Delta v_p \\
&2)\ \hat{v}_{i,\,x}(K)-\hat{v}_{i-1,\,x}(K)= 0\ , \qquad \text{if } i \in \mathbf{G} \text{ and } i \ge 2 \\
&3)\ \left|\hat{v}_{i,\,x}(K)-\hat{v}_{i-1,x}(K)\right| \le \Delta v_c \quad \text{if } i \in \tilde{\mathbf{G}} \text{ and } i \ge 2 \\
&4)\ \left|\hat{v}_{i,\,x}(K)-\hat{v}_{i,\,x}(K-1)\right| \le \Delta v_c, \quad \forall K \ge 2 \\
&5)\ \hat{v}_{i,\,x}(K) \in V_L, \\
&6)\ \hat{B}^{\bar{x}}_{i,\,x}(K) \cdot \hat{B}^{x}_{i,\,\bar{x}}(K)=0, \\
&7)\ \hat{B}^{\bar{x}}_{i,\,x}(K) - \hat{B}^{\bar{x}}_{i-1,\,x}(K) = 0, \qquad \forall i \ge 2 \\
&8)\ \sum_{K=2}^{K_N}\left|\hat{B}^{\bar{x}}_{i,\,x}(K) - \hat{B}^{\bar{x}}_{i,\,x}(K-1)\right| \le N_B, \quad \forall i \ge 2 \\
&9)\ 0 \le \hat{p}^{i+1,\bar{x}}_{i,\,x,\,\mathrm{CA}}(K) \le 1, \qquad \forall i \le N-1 \\
&10)\ \hat{p}^{i+1,\,\bar{x}}_{i,\,x,\,\mathrm{CA}}(K) \cdot \hat{p}^{i+1,\,x}_{i,\,\bar{x},\,\mathrm{CA}}(K)=0. \\
&11)\ 0 \le \hat{L}(K) \le T_c \\
&12)\ \hat{L}(K) \in N
\end{aligned}
\tag{24}
$$

Detailed interpretations of (24) are as follows. First, at the same longitudinal position, vehicles (both RHVs and CAVs) on parallel lanes are compelled to follow speed limits without sharp variation. Because implementing sharply varied speed limits might encourage a large quantity of DLC maneuvers on lateral dimension, which would further reduce the freeway capacity and induce instability in traffic flow, especially in the case of a traffic incident or over-saturated traffic conditions. As stated by the 1^{st} constrain of (24), the maximal VSL variation Δv_p of two adjacent lanes is restricted as a small value.

As depicted in Figure 5, G_V denotes a freeway segment governed by VMS V, whereas $\tilde{G}_V$ denotes the boundary cell of this segment connecting the segment governed by VMS $V-1$. Let set $\mathbf{G}$ to be the collection of all G_V, for $V= 1, 2, \ldots, N_V$, where N_V is the number of VMS gantries installed along the study site. Similarly, $\tilde{\mathbf{G}}$ denotes the combination of all boundaries $\tilde{G}_V$. Taking into account the operational feasibility and comfort in driving, the 2^{nd} and 3^{rd} items of (24) constrain the VSL fluctuation on the longitudinal dimension. The 2^{nd} equality suggests that the VSL should be the same within the segment governed by a gantry. As frequent variation of VSL from cell to cell would trigger oscillations in the traffic flow, and further induces instability, typically the cell length is usually not very long, e.g., 0.25 mile in numerical example. Furthermore, the VMS gantries cannot be installed so densely as to disseminate cell-to-cell VSL control to RHVs. Even if gantries are installed for each cell, human drivers cannot adapt to such quickly varying speed limit control.

The 3^{rd} inequality suggests that the VSLRs which are simultaneously released by two adjacent VMS gantries should not admit sharp variation. The term Δv_c denotes the maximal tolerable fluctuation of VSL between two adjacent boundary cells which are governed by two successive VMS segments on the longitudinal dimension, such as between $\tilde{G}_V$ and G_{V-1}. In this study, the value of Δv_c is restricted to ensure that the new speed limit could be achieved within a simulation step, e.g., $\Delta v_c = 20$ miles/hour can be achieved in 10 seconds by acceleration/deceleration. The same fluctuation restriction Δv_c is also applied to constrain the temporal variation of VSL issued by each VMS, as expressed in the 4^{th} inequality of (24).

Also, the increment (or decrement) (also termed as quantization in Zhang and Ioannou (2017)) of VSL must be practical for human drivers. For example, a change of 1 mile/hour in the VSL might be imperceptible to human drivers, whereas a change of 5 miles/hour is quite perceptible to drivers and thus may be more suitable. To this end, in this paper, the speed limit control is discretized as the finite value set $V_L = \{10,\ 15,\ \cdots\ v_f - 5,\ v_f\}$, with increments of 5 miles/hour , saying the lowest VSL issued by the system is 10 miles/hour, and the upper bound of VSL is set to be the free flow speed v_f under normal conditions, i.e., 70 mile/hour. The five items as introduced above constrain the variables associated with the VSLR released by the VMS, while the VSLC issued by the VACS follows the same constraints.

Some of the constraints have been included in different forms in the literature. For example, Zhang and Ioannou (2017) considered the discretization, quantization of VSL and saturation of speed limit variations. The constraint set proposed here is more general than that in Zhang and Ioannou (2017) since both spatial-temporal saturation of speed limit variations meanwhile asynchronous simulation steps for traffic flow model and the cycle for control implementation are considered in this paper.

The 6^{th} formula of (24) suggests that the opposite lane changing orders cannot be released on the same VMS segment during the same control cycle, and 10^{th} equation suggests the similar constrains for CAVs, for avoiding conflicts proposed by control measures. The 7^{th} equation ensures the consistency of LCRs issued by the VMSs upstream to the target turning location, e.g., an incident spot. The 8^{th} inequality prevents the case that the RHVs are suggested to keep changing their lanes more than N_B as frequent lane-changing maneuvers would disrupt traffic stability, where N_B is set to be 2 in the numerical example this paper. Finally, the duration of red phase of the ramp metering should be a integer between 0 and control cycle time T_c, i.e., the 11^{th} and 12^{th} constraints.

To provide a general overview, a flow chart of implementing the multiclass multilane traffic model for optimal control design is summarized in Figure 6, wherein the symbols in blue color correspond to the control variables.

4. A model-based reinforcement learning algorithm

4.1. A conceptional structure of the model-based reinforcement learning algorithm

The optimal control problem for general nonlinear systems usually adopts gradient-based approaches to devise solution algorithms based on several assumptions, such as the system dynamics (governing the state space) is at least continuously differentiable whereas the control is continuous. It can be seen from the development of the multiclass multilane traffic flow model that the underlying system dynamics of our optimal control problem is not differentiable, and it is too complicated to evaluate the gradient of the objective function along the state trajectory and the constraint set. As a matter of fact, even its simplest case, i.e., the CTM case, is non-differentiable (Zhong et al., 2014, 2016b). On the other hand, as explained in the previous section, the control is also subject to several constraints for practical considerations, e.g., the VSL is discretized, and the control is implemented in a cycle-to-cycle manner (2 min interval while the simulation time step is 10 s). To this end, we will not solve the optimal control directly but try some heuristic optimization algorithms adopted from reinforcement learning (RL) on the equivalent reformulation of the optimal control problem.

A conceptual overview of the overall training loop for the proposed model-based reinforcement learning algorithm is depicted in Figure 7. Under this framework, we have a known model (i.e., the proposed multiclass multilane CTM) and use planning (i.e., the dynamic optimization) to learn a global value (i.e., the objective function) and policy (i.e., the optimal control). A famous example adopting such kind of framework is AlphaZero (Silver et al., 2017). Model-based RL methods eventually seek a global approximation of the optimal value and/or policy function. The planning result may be used to update this global approximation. The arrows indicate possible interactions between different components. First, the model (Arrow a) feeds the planning component with the dynamic traffic state to optimize freeway performance. It directs new planning iterations based on learned knowledge in value and/or policy functions (Arrow b). In the meantime, it uses the planning output to update learned value and/or policy functions (Arrow c). In practice, the learned optimal control is implemented to select actions in the real world and generate data (Arrows d & e). The data generated are observations of the effect of the control implemented. The data are then used to update/recalibrate the model and update the learned value and/or policy functions (Arrows f & g) to form a closed-loop. However, we need to point out that this closed-loop has not yet formed in this paper since there is no real-world

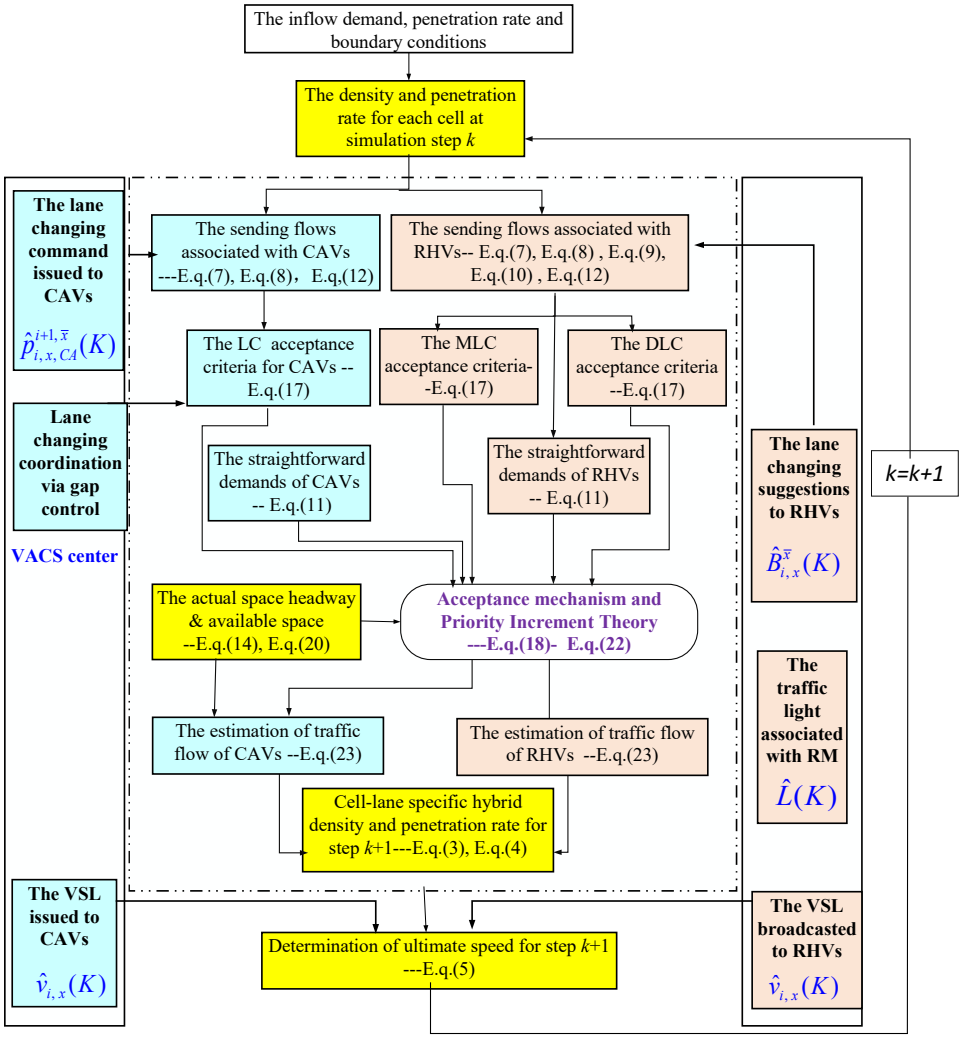


Figure 6: A flowchart of multiclass multilane traffic model as the network loading model for optimal control

test field with sufficient CAVs running on the freeway and thus no real-data for training and update. Therefore, we put our priori trust in the known model.

Gradient-based planning approaches have been widely adopted to devise solution algorithms in reinforcement learning under several common assumptions. For instance, the gradient-based approach requires a differentiable model, i.e., the system dynamics (governing the state space) is assumed at least continuously differentiable, whereas the control is continuous. It can be seen from the development of the multiclass multilane traffic flow model that the underlying system dynamics of the optimal control problem concerned is, however, not differentiable. Therefore, it is too complicated to evaluate the gradient of the objective function along the state trajectory and the constraint set. To this end, we will not solve the optimal control directly but attempt to use a gradient-free reinforcement learning algorithm, i.e., a Cross-Entropy-Method-based (CEM-based) optimal control search algorithm.

Control discretization heavily affects the solution quality obtained by the RL and ADP. Fine control discretization, which is generally difficult to be designed, is an essential prerequisite. Rather than searching the optimal control directly, we use control parameterization to parameterize the control (or policy) and search for the optimal parameters that lead to maximal returns (or minimizing the objective function) to reduce the dimension. Control (or policy) parameterization technique is a popular scheme for solving the optimal control problem using RL and ADP. For example, neural networks and the radial basis functions (RBFs) are widely used to approximate the control or action space to

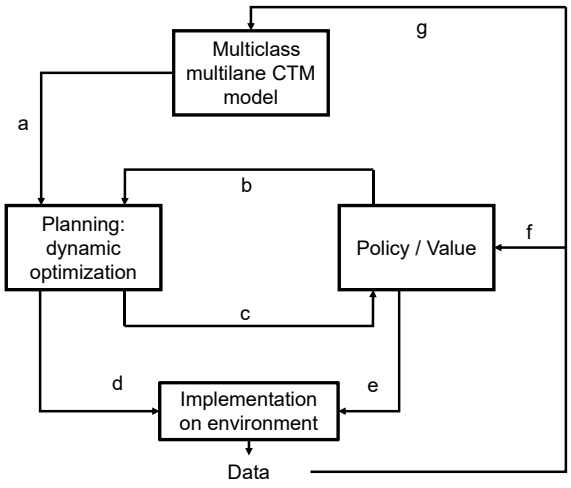


Figure 7: A conceptional structure of the model-based reinforcement learning algorithm

reduce dimension. By the universal approximation theorem, we can approximate the control[6] by a linear combination of basis functions (or a vector of features). By this, the coefficients in the linear combination are decision variables to be determined rather than solving the optimal control problem directly. Applying the control parameterization technique yields a finite-dimensional approximation of the original optimal control problem which is of infinite-dimension (if it is in continuous-time) or very high dimension DP problem (if it is in discrete-time) and is generally difficult to solve. Control parameterization technique yields a low dimensional parameter optimization approximation of the original high dimensional optimal control problem. By such parameter optimization approximation, we can directly use the CEM optimization process for parameter optimization developed in Zhong et al. (2016a) to yield an efficient computation algorithm for the optimal control problem.

4.2. Control parameterization

To begin with, we present the discretization of the control (or the actions). As previously discussed, the VSL control is discretized into finite values in 5 miles/hour increments, e.g., $[10, 15, \cdots, 60, 65, 70]$ miles/hour for several practical issues. Indeed the lane flow distribution ratio needs to be discretized similarly. Taking a cell with a length of 1/8 miles as an example, if we assume that the density is 160 veh/mile/lane (very congested traffic), the number of vehicles in the cell is 20. An increment of 5% is sufficient (one vehicle) for the lane flow distribution ratio. All the elements from the discretization of an originally larger (e.g., continuous) action space U consist of the new action space of the problem. We assume that this set contains M distinct actions and is denoted as $U_d = \{u_1, \cdots, u_M\}$. We further assume the state space X comprises D variables.

Towards this, to enable a random search by the CEM and to cover the effects missed by the control discretization, we present a control parameterization in line with that of Busoniu et al. (2010). We use $\aleph$ basis functions (BFs), i.e., $\varphi_i(x;\xi) : X \to \mathbb{R}$, $i = 1, \cdots, \aleph$, defined over the state space and parameterized by a vector ξ that gives the locations and shapes of the BFs. The BFs are associated with (control) actions by a many-to-one mapping, which can be represented as a vector $\vartheta \in \{1, \cdots, M\}^{\aleph}$ that associates each BF φ_i to a discrete (control) action index ϑ_i, or equivalently to a discrete action u_{ϑ_i}.

As discussed in Busoniu et al. (2010), the number $\aleph$ of BFs, in conjunction with the type of BFs, determine the accuracy of the control (or policy) approximation. Given the type of BFs, a good value of $\aleph$ for a given problem cannot be determined theoretically in general but could be found empirically. In the literature, the use of a control parameterization technique for optimal control of a relatively small number of BFs is often sufficient to provide a

[6] Any vector in the functional space can be represented by a linear combination of its basis functions.

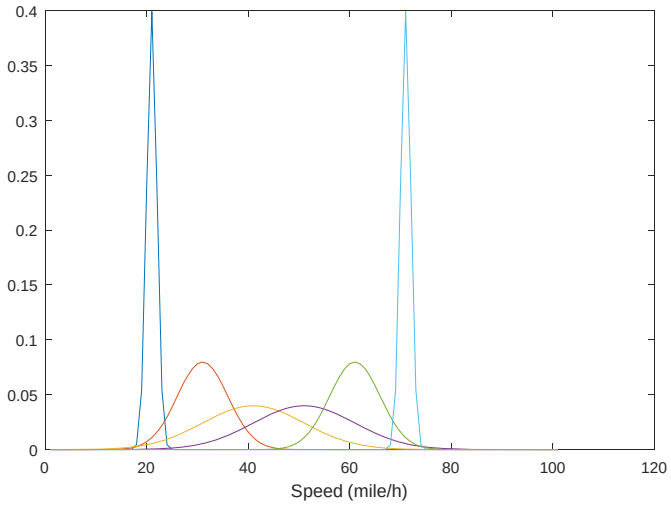


Figure 8: Illustration of the policy parameterization by radial basis functions

good policy approximation. If prior knowledge about the complexity of an optimal policy is available, one can choose a reasonable type of BF and the value of $\aleph$ beforehand such as in our case.

There are many types of BFs, such as splines, polynomials, and RBFs, which could be used. In this paper, we chose the Gaussian RBFs to adapt to the Gaussian assumption that we adopted in the CEM sampling process. The Gaussian RBFs are defined by

$$\varphi_i(x;\xi) = \exp\left(-\sum_{d=1}^{D}\left(\frac{(x_d - c_{i,d})^2}{b_{i,d}^2}\right)\right) \tag{25}$$

where D is the number of state variables, $c_i = [c_{i,1}, \cdots, c_{i,D}]^T$ is the D-dimensional center of the i^{th} RBF, and $b_i = [b_{i,1}, \cdots, b_{i,D}]^T$ is its width. Denoting the vector of centers by $\boldsymbol{c} = [c_1^T, c_2^T, \cdots, c_{\aleph}^T]^T$ and the vector of widths by $\boldsymbol{b} = [b_1^T, b_2^T, \cdots, b_{\aleph}^T]^T$, the BFs parameter vector is $\xi = [\boldsymbol{c}^T, \boldsymbol{b}^T]^T$. In our case, all the elements of the vectors of center and width are non-negative.

Using the VSL control as an example, a schematic representation of the parameterization is illustrated in Figure 8. The spike of the RBF is determined by the parameters of the basis function. For example, because 20 miles/h lies at the boundary of the speed limit control, one may use a spiked RBF to parameterize this control, whereas one may use an RBF with a high standard deviation to parameterize the VSL controls in the middle (which may yield a many-to-one mapping). For any state x, the control associated with a BF that takes the largest value at x will be chosen (Busoniu et al., 2010), i.e., $h(x;\xi,\vartheta) = u_{\vartheta_{i^*}}$, where $i^* = \arg\max_i \varphi_i(x;\xi)$.

4.3. Recapitulation of the cross-entropy method for optimization

Due to the complexity of the multiclass multilane traffic flow model as previously discussed, the optimization problem is with the non-differentiable complicated functional structure of the control vector and the constraint set, which would induce many local optima. Gradient-based policy search (or iteration) approaches may not be good choices. A gradient-free approach, i.e., a Cross-Entropy-Method-based (CEM-based) algorithm adopted from the reinforcement learning community for learning the game of Tetris (Szita and Lorincz, 2006) was extended to solve the problem concerned. Rather than searching the optimal control directly, this algorithm searches for the optimal parameters (yielded from control parameterization) that lead to maximal returns (or minimizing the cost). Thiery and Scherrer (2009) optimized the weights with the CEM that led to one of the best publicly known controllers to the game of Tetris (Gabillon et al., 2013). Note that, after the control parametrization, the decision variables (i.e., the parameters) are continuous. The CEM-based algorithm proposed in this paper is extended from Zhong et al. (2016a) for continuous optimization problem in transportation.

The Kullback-Leibler (K-L) divergence, denoted as $D_{KL}(p||q)$, is used to measure the dissimilarity of two probability distributions, e.g., $p(x)$ and $q(x)$. For discrete probability distributions p_j and q_j, the K–L divergence of q from p is defined as

$$D_{\mathrm{KL}}(p||q) = \sum_j p_j \ln \frac{p_j}{q_j}.$$

For continuous distributions $p(x)$ and $q(x)$, the K–L divergence is defined to be

$$D_{\mathrm{KL}}(p||q) = \int_{-\infty}^{\infty} p(x) \ln \frac{p(x)}{q(x)} \, \mathrm{d}x.$$

To begin with, we provide a recapitulation of the CEM for general optimization problem following Rubinstein and Kroese (2004). Consider a general minimization problem of the following form:

$$\gamma^* = \min_{x \in \chi} S(x), \tag{26}$$

where γ^* represents the (global) minimum of $S(x)$ and x is defined in a certain function space χ. Obtaining an (global) optimum solution for a general optimization problem can be regarded as a rare event, especially when the problem is nondifferentiable and nonconvex. To be specific, a rare event such as $S(x) \leq \gamma$ is defined for the minimization of $S(x)$ by defining the minimum γ^* as a threshold (or some $\gamma \geq \gamma^*$ but sufficiently close to γ^*). The CEM is used to formulate a family of probability density functions (PDFs) distributed in χ, denoted by $f(x;v)$, parameterized by v to apply the Monte Carlo approach to estimate above rare event probability. By the above rare event analogy, we can define

$$\ell(\gamma) = P_u(S(X) \leq \gamma) = E_u\left(I_{\{S(X) \leq \gamma\}}\right), \tag{27}$$

where $X = (X_1, X_2, \cdots, X_n)$ is a random vector generated by PDF with the parameter v set to u (i.e., u is a realization of v) in $f(x;v)$. P_u denotes the evaluation of probability. E_u denotes the expectation. $I(\cdot)$ is the indicator function, i.e., $I_{\{S(X) \leq \gamma\}} = 1$, if and only if $S(X) \leq \gamma$ is true, 0 otherwise. The original optimization problem is then converted into a rare event probability estimation problem by the CEM, and the objective is to maximize γ such that $\ell(\gamma)$ approaches 0. In ther words, γ is the minimum of $S(x)$ in a probabilistic sense because γ gives the largest value such that $S(X)$ is greater than γ, i.e., $S(x) > \gamma$ with very high probability if $\ell(\gamma)$ approaches 0. For a given γ, a practical way to estimate $\ell(\gamma)$ is to generate some samples from χ and perform Monte Carlo simulations. To this end, N samples are generated from $f(x;u)$ and $\ell(\gamma)$ is estimated as

$$\widehat{\ell} = \frac{1}{N} \sum_{i=1}^{N} I_{\{S(\mathbf{x_i}) \leq \gamma\}},$$

where $\mathbf{x_1}, \cdots, \mathbf{x_N}$ denotes a random sample. This crude Monte Carlo simulation idea is computational intensive in that it requires a huge number of samples to accurately estimate $\ell(\gamma)$ when $S(X) \leq \gamma$ is a rare event. To overcome the computation problem, the CEM exploits the power of the importance sampling technique, which uses a different probability density function, $\kappa(x;\epsilon)$, on χ and computes the estimation of $\ell(\gamma)$ as $\widehat{\ell}(\gamma)$ by

$$\widehat{\ell}(\gamma) = \frac{1}{N} \sum_{i=1}^{N} I_{\{S(\mathbf{x_i}) \leq \gamma\}} \frac{f(\mathbf{x_i};u)}{\kappa(\mathbf{x_i};\epsilon)}.$$

Defining

$$\kappa^*(x,\epsilon) = \frac{I_{\{S(x) \leq \gamma\}} f(x;u)}{\ell(\gamma)},$$

and replacing κ by κ^*, we have

$$\widehat{\ell}(\gamma) = \frac{1}{N} \sum_{i=1}^{N} I_{\{S(\mathbf{x_i}) \leq \gamma\}} \frac{f(\mathbf{x_i};u)}{\kappa^*(\mathbf{x_i};\epsilon)} = \ell(\gamma).$$

The problem now turns to the estimation of κ^* because $\ell(\gamma)$ is unknown. The CEM defines the distance between two

PDFs, $\kappa(x;\epsilon)$ and $f(x;v)$, using K-L divergence:

$$D_{KL}(\kappa\|f) = \int \kappa(x;\epsilon)\ln\kappa(x;\epsilon)dx - \int \kappa(x;\epsilon)\ln f(x;v)dx,$$

where E_κ means the expectation under the probability density function κ, and $D_{KL}(\kappa\|f) \geq 0$ with $D_{KL}(\kappa\|f) = 0$ if and only if $f(x) = \kappa(x)$. Minimizing the K–L distance to approach κ^* is equivalent to

$$\max_v \int \kappa^*(x,\epsilon)\ln f(x;v)\,dx,$$

by selecting v.

Substituting the definition of κ^*, we arrive at

$$\max_v \int \frac{I_{\{S(x)\leq\gamma\}}f(x;u)}{\ell(\gamma)}\ln f(x;v)dx.$$

Using the definition of $\ell(\gamma)$, an equivalent optimization problem is

$$v^* = \arg\max_v E_u I_{\{S(X)\leq\gamma\}}\ln f(X;v).$$

Note that the nominal PDF f is assumed to be parameterized by a finite-dimensional vector u, i.e., $f(x) = f(x;u)$, and that the importance sampling PDF is $f(\cdot;v)$ for some parameter v. Again, using importance sampling, with a change of measure $f(\cdot;w)$ the above optimization problem is rewritten as

$$v^* = \arg\max_v E_w I_{\{S(X)\leq\gamma\}}W(X;u,w)\ln f(X;v),$$

for any reference parameter w, where

$$W(x;u,w) = \frac{f(x;u)}{f(x;w)},$$

is the likelihood ratio at x between $f(\cdot;u)$ and $f(\cdot;w)$ (Rubinstein and Kroese, 2004). Finally, v^* is estimated by solving the following stochastic program:

$$\hat{v}^* = \arg\max_v \frac{1}{N}\sum_{i=1}^{N} I_{\{S(\mathbf{x_i})\leq\gamma\}}W(\mathbf{x_i};u,w)\ln f(\mathbf{x_i};v) \doteq \arg\max_v \hat{D}(v),$$

where $\mathbf{x_1},\cdots,\mathbf{x_N}$ is a random sample from $f(\cdot;w)$. In typical applications the function $\hat{D}$ is convex and differentiable with respect to v. The solution may be readily obtained by solving the following system of equations (with respect to v).

$$\frac{1}{N}\sum_{i=1}^{N} I_{\{S(\mathbf{x_i})\leq\gamma\}}W(\mathbf{x_i};u,w)\nabla_v\ln f(\mathbf{x_i};v) = 0,$$

where ∇_v indicates the gradient with respect to v.

To sum up the above description briefly, the CEM comprises two key steps:

1. Generate trial decision variable sets randomly according to the chosen distributions, which are usually specified as Bernoulli or Uniform distributions (for discrete variable) or Gaussian distribution (for continuous variable).

2. Update the probability distribution used to generate the random trial sets according to the principle of “importance sampling".

Table 4: Notation list of the CEM for parameter optimization

Notation	Description
N	Number of samples
$\hat{J}(x_0, \xi, \vartheta)$	Objective function
ξ, ϑ	Decision parameters
$\mathrm{N}(\cdot; v)$	Normal distribution of model parameters with parameters v
$v_{\xi_0} = [\mu_{\xi_0}, \delta_{\xi_0}]$	Parameters of density with mean and standard deviation
ρ	Percentage of elite samples on all samples
α	Smoothing parameter
ε	Prescribed threshold
X_{im}, X_{un}	The important/unimportant parameter combinations
δ	Bandwidth
w	The flag for removing unimportant parameter combinations
ε_1	Threshold of adding unimportant parameter combinations

4.4. Optimal parameter search based on the CEM

Through the control parameterization depicted in Section 4.2, the optimal control problem is converted to a parameter optimization problem. From the CEM optimization process for parameter optimization as depicted in the above section, we need to search the mean and variance for each decision variable. Therefore, the density for each center $c_{i,d}$ is parameterized by its mean $\mu_{c_{i,d}}$ and its standard deviation $\delta_{c_{i,d}}$, whereas the PDF for a width $b_{i,d}$ is likewise parameterized by $\mu_{b_{i,d}}$ and $\delta_{b_{i,d}}$. The vector of the RBF parameters is $\boldsymbol{v}_\xi = [(\boldsymbol{\mu_c})^T, (\boldsymbol{\delta_c})^T, (\boldsymbol{\mu_b})^T, (\boldsymbol{\delta_b})^T]^T$. If the centers are chosen beforehand, e.g., the centers of the RBFs associated with the VSL are specified as [10, 15, 20,$\cdots$,60, 65, 70] miles/hour, then the vector of the centers is not a decision one and $\boldsymbol{v}_\xi = [(\boldsymbol{\mu_b})^T, (\boldsymbol{\delta_b})^T]^T$. The dimension of the problem can then be reduced but at the price that the control law may be a Pareto optimum.

The original dynamics are now approximated by incorporating the parameterized control as input.

$$x_{k+1} \approx F(x_k, h(x; \xi, \vartheta)) = \hat{F}(x_k, \xi, \vartheta).$$

The function $F(x_k, .)$ is only a compact representation of the multiclass multilane traffic flow model, whereas $\hat{F}(x_k, \xi, \vartheta)$ is its approximation by incorporating the parameterized control as input. The original cost function is approximated by

$$J(x_k) \approx \hat{J}(x_k, \xi, \vartheta)$$

We use this form to highlight that a) the dynamics originally driven by the control input is now determined by the parameters of the parameterized control so does the cost function; b) finding an optimal control is now about searching optimal parameters to minimize the approximated cost function. Unlike the study (Busoniu et al., 2010), the multiclass multilane traffic flow model is a deterministic process. It suffices to simulate a single trajectory initialized by x_0 in X_0 to find the optimal control sequence associated with x_0. To this end, we define the score function (i.e., the objective function to be minimized) as

$$\hat{J}(x_0, \xi, \vartheta) = \sum_{k=1}^{k_N} \hat{J}(x_k, \xi, \vartheta)$$

Now, we are ready to describe the CEM for policy approximation as summarized in Algorithm 4.4 below.

5. Numerical simulation

5.1. Description of the "test site"

A numerical study is conducted on a virtual 2.75-mile-long freeway segment with two lanes, as shown in Figure 9. It is assumed that an on-ramp is connected to this freeway segment near the downstream boundary. Under normal

ALGORITHM
The cross-entropy method for parameter optimization

Input: Scoring function $\hat{J}(x_0, \xi, \vartheta)$, density function $N(\cdot; v)$, other parameters
1: initialize density parameters $v_{\xi_0} = [\mu_{\xi_0}, \delta_{\xi_0}]$, and $v_{\vartheta_0} = [\mu_{\vartheta_0}, \delta_{\vartheta_0}]$
2.1: generate original samples $\xi_O = [\xi_1, \cdots, \xi_{\mathbf{N}}]$ from $N(\xi; v_{\xi_0})$ and $\vartheta_O = [\vartheta_1, \cdots, \vartheta_{\mathbf{N}}]$ from $N(\vartheta; v_{\vartheta_0})$, and $Sam_O = [\xi_O, \vartheta_O]$
2.2: fix one parameter ξ^i or ϑ^j to produce derived samples $Sam_O|\xi^i$ *or* ϑ^j with $i = 1, ..., d$
2.3: compute $S(x) = \hat{J}(x_0, \xi, \vartheta)$, and use kernel density method to estimate original PDF and derived PDF: $\hat{g}(S(x)) = \frac{1}{N\delta} \sum_{i \in I} \frac{1}{2\pi}(-\frac{1}{2}(\frac{S(x) - \hat{S}(x_i)}{\delta})^2)$
2.4: calculate the cross-entropy distance between original PDF $f(x)$ and derived PDF $g(x|X_i)$, and determine the important parameter combinations $X_{im}$and unimportant parameter combinations X_{un}
3: $t \leftarrow 0$ and $w \leftarrow 0$, fix X_{un} to its constant value, remove v_{un} from v_{ξ_0} and v_{ϑ_0} respectively, and initialize x_0
4: repeat
5: $t \leftarrow t + 1$
6: if $\max\{v_{\xi_{t-1}}, v_{\vartheta_{t-1}}\} \leq \varepsilon_1$ and w=0
7: add $v_{un,0}$ into v_{ξ_0} and v_{ϑ_0} respectively, and set $w = 1$
8.1: generate samples $\xi_1, \cdots, \xi_{\mathbf{N}}$ from $N(\xi; v_{\xi_{t-1}})$ and $\vartheta_1, \cdots, \vartheta_{\mathbf{N}}$ from $N(\vartheta; v_{\vartheta_{t-1}})$
8.2: compute $\hat{J}(x_0, \xi_i, \vartheta_i)$ with i = 1,. . .,**N**
8.3: reorder and reindex s. t. $\hat{J}_0 \leq \cdots \leq \hat{J}_{\mathbf{N}}$
8.4: $\gamma_t \leftarrow \hat{J}_{\lceil (1-\rho)\mathbf{N} \rceil}$
8.5: $\hat{v}_{\xi_t} \leftarrow [\mu_{\xi_t}, \delta_{\xi_t}]$ where $u_{\xi_t} = \frac{1}{\rho N} \sum_{j=1}^{\rho N} \xi_{\lceil j \rceil}$, $\delta^2{}_{\xi_t} = \frac{1}{\rho N} \sum_{j=1}^{\rho N} (\xi_{\lceil j \rceil} - \mu_{\xi_t})^2$
$\hat{v}_{\vartheta_t} \leftarrow [\mu_{\vartheta_t}, \delta_{\vartheta_t}]$ where $u_{\vartheta_t} = \frac{1}{\rho N} \sum_{j=1}^{\rho N} \vartheta_{\lceil j \rceil}$, $\delta^2{}_{\vartheta_t} = \frac{1}{\rho N} \sum_{j=1}^{\rho N} (\vartheta_{\lceil j \rceil} - \mu_{\vartheta_t})^2$
8.6: $v_{\xi_t} = \alpha \hat{v}_{\xi_t} + (1 - \alpha) v_{\xi_{t-1}}$ and $v_{\vartheta_t} = \alpha \hat{v}_{\vartheta_t} + (1 - \alpha) v_{\vartheta_{t-1}}$
9: until $\max\{\delta_{\xi_t}, \delta_{\vartheta_t}\} \leq \varepsilon$ or $t = t_{\max}$
Output: μ_{ξ_t} and μ_{ϑ_t}, the best sample; and γ_t, the best score

traffic conditions, the only lateral movement on the upstream of this segment is the DLC for achieving a speed advantage. During the rush hours, a temporary bottleneck would form due to the weaving behavior. An integration of ramp metering, lane changing control, and variable speed limit control is imposed to alleviate the traffic congestion. The communication infrastructures that transmit real-time traffic data and issue lane-changing control (LCC) and variable speed limit controls (VSLC) to CAVs are installed along the freeway section. It is also assumed that there are three VMS gantries, namely VMS A, VMS B and VMS C, installed along the segment that broadcasts lane-changing recommendation (LCR) and variable speed limit recommendation (VSLR) to RHVs. The simulation time step is 10 s. To meet the numerical stability requirement, i.e., the Courant-Friedrichs-Lewy (CFL) condition of the CTM, the cell length should be greater than or equal to the distance traveled during a simulation time step at the free-flow speed. Therefore, this 2.75-mile-long freeway segment is divided into 11 cell packages with equal lengths of 0.25 miles while the on-ramp is assumed to be located at the boundary of cell 10 and cell 11.

The initial simulation time is 0. To maintain the same total throughput, the inflow demand is set to be 0 in the last 5 minutes. Lane-specific inflow demand function $S_i(k)$ (P.C.U./hour/lane) including both CAV and RHV is assumed:

$$S_i(k) = \begin{cases} s, & 0 \leq k \leq 40 \text{ min}; \\ 0, & k > 40 \text{ min}. \end{cases} \tag{28}$$

The on-ramp demand function $\varpi(k)$ (P.C.U./hour) is assumed:

$$\varpi(k) = \begin{cases} 0, & 0 \leq k < 5 \text{ min}; \\ q_n, & 5 \leq k \leq 25 \text{ min}; \\ 0, & k > 25 \text{ min}. \end{cases} \tag{29}$$

The values of s and q_n vary with respect to the penetration rate to assess the effectiveness of the optimization framework under different scenarios as outlined in Table 5. For simplicity, we assume both $S_i(k)$ and $\varpi(k)$ share the same

Table 5: Settings of the three test scenarios

Scenario ♯	*PN*	Nominal Capacity under this *PN* (P.C.U./hour)	mainstream demand s (P.C.U./hour/lane)	On-ramp demand q_n (P.C.U./hour)	Cumulative inflow for the whole horizon (P.C.U.)
Scenario 1	66.7%	3301	2700	600	3800
Scenario 2	33.3%	2294	1800	399	2533
Scenario 3	0	1719	1350	300	1900

penetration rate (*PN*) for each scenario.

The driving behavior of RHVs is affected by the information on the VMSs when they pass the gantry, e.g., the drivers will not notice the information on the VMS B until they arrive at Cell 5. As expected, the compliance to these recommendations will gradually affect the movements of RHVs for downstream cells until they receive new information from the downstream gantry. *Please note that the CAVs are always with full automatic control, otherwise they are regarded as RHVs. Therefore, when we mention "without control" we mean that the CAVs operate under automated vehicle mode (i.e., single vehicle intelligence) rather than to coordinate the movements of the RHVs and CAVs simultaneously.*

As discussed in Sections 2-4, considering the applicability, the VMS information should not vary too frequently nor too sharply on temporal or spatial domains; otherwise, drivers will have difficulty in making decision and taking action. In this example, the control cycle T_c is set to be 120 sec, which consists of 12 simulation steps (with duration time of each simulation step T_s = 10 sec). Within each control cycle, the control variables $\hat{v}_{i,x}(K)$, $\hat{B}^{\bar{x}}_{i,x}(K)$, $\hat{p}^{i+1,\bar{x}}_{i,x,\mathrm{CA}}(K)$ and $\hat{L}(K)$ maintain the same values, while during the whole simulation horizon, their variations are subject to the constraints outlined in (24).

For determining the time-dependent fundamental diagram and minimum space headway acceptance criteria, the response time is set to ΔT_{RH} =1.85 sec for RHVs and ΔT_{CA} =0.35 sec for CAVs. Note that the response time of CAVs is assumed to vary from 0.25 sec to 1.5 sec in Levin and Boyles (2016a). The minimum headway required by CAVs is about 0.6 sec if the response time is chosen as 0.35 sec. Since the minimum headway required by CAVs would heavily affect the capacity of the freeway with mixed traffic of RHVs and CAVs, we will conduct a sensitivity analysis on several typical response time values of CAVs. All vehicles are assumed to be P.C.U. 20 ft (6 meters) long, and the safe constant gap is 6.5 ft (2 meters, see the empirical study in Pan et al. (2016)). It is assumed that the compulsory speed limit is 70 miles/hour under normal conditions. The parameters for evaluating the distribution of MLC and DLC demands (i.e., Equations (9)-(10)) are: $\alpha_1 = 671$, $\alpha_2 = 33.7$, $d_c = 0.05$ miles, $\tau = 3$ sec. These parameters were calibrated and validated in Pan et al. (2016). The maximal tolerable variation of VSL between successive control cycles (restriction on temporal dimension) and successive VMS gantries (restriction on longitudinal dimension) is Δv_c = 20 miles/hour, and the maximal tolerable variation between parallel lanes (restriction on lateral dimension) is Δv_p = 20 miles/hour. Besides, the LCR will disseminate not greater than 2 lane-changing recommendations to a vehicle when traversing the test section, i.e., $N_B = 2$.

In this numerical study, the total delay is converted to monetary cost so that the fuel price, emission cost, and electricity cost can be included as well. Following the settings in Tang et al. (2017), the value of travel time is V_t = 24 USD/hour. The cost of fuel c_f = 0.64 USD/L. The emission cost of CO and HC is $c_{\mathrm{CO}} = c_{\mathrm{HC}}$ = 0.07 USD/mg, while that of NO is c_{NO} = 0.007 USD/mg. The electricity cost is c_c = 0.12 USD/kWh. The speed limit should be not less than 10 miles/hour, sayingV_L = {10, 15, $\cdots$ 65, 70} . The fuel consumption matrix K_f, the electricity cost matrix K_c, and the emission cost matrices K_{NO}, K_{CO}, K_{HC} are:

$$K_f = \begin{bmatrix} -0.679439 & 0.135273 & 0.015946 & -0.001189 \\ 0.029665 & 0.004808 & -2.0535e^{-5} & 5.5409285e^{-8} \\ -0.000276 & 8.3329e^{-5} & 9.37e^{-7} & -2.479644e^{-8} \\ 1.487e^{-6} & -6.1321e^{-5} & 3.04e^{-7} & -4.467234e^{-9} \end{bmatrix}$$

$$K_c = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 2.28 & 0 & 0 \\ 1.8e^{-4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad K_{CO} = \begin{bmatrix} 0.887447 & 0.148841 & 0.03055 & -0.001348 \\ 0.0710 & 0.00387 & 9.32e^{-5} & -7.06e^{-7} \\ -7.86e^{-4} & 9.26e^{-5} & 4.92e^{-5} & -3.14e^{-7} \\ 4.62e^{-6} & 4.6e^{-5} & -1.41e^{-6} & 8.17e^{-9} \end{bmatrix}$$

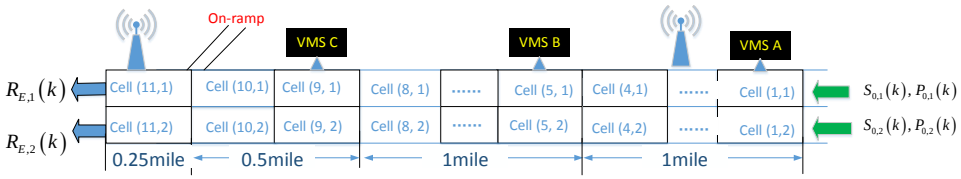


Figure 9: Topology of the “test site"

$$K_{HC} = \begin{bmatrix} -0.728042 & 0.012211 & 0.023371 & -9.3243e^{-5} \\ 0.024950 & 0.010145 & -1.03e^{-4} & 6.18e^{-7} \\ -2.05e^{-4} & -5.49e^{-4} & 3.7592e^{-5} & -2.13e^{-7} \\ 1.949e^{-6} & -6.1321e^{-5} & 3.31e^{-6} & -1.739372e^{-8} \end{bmatrix}$$

$$K_{NO} = \begin{bmatrix} -1.067682 & 0.254363 & 0.008866 & -9.51e^{-4} \\ 0.046423 & 0.015482 & 5.69e^{-7} & 3.28e^{-7} \\ -1.73e^{-4} & 0.002876 & -5.866e^{-5} & 2.4e^{-7} \\ 5.69e^{-7} & -0.000321 & 1.943e^{-6} & -1.257413e^{-8} \end{bmatrix}$$

For all scenarios, simulations are conducted for the following 2 cases:

- **Case 1: Baseline (without optimal control)**. No optimal control or recommendations are disseminated to CAVs or RHVs for coordinating their movements. CAVs might execute MLC maneuvers by approaching to the bottleneck, and RHVs might propose DLC maneuvers when speed advantage is observed on the parallel lane.
- **Case 2: Integrated optimal control strategies**. LCC, VSlC and RM (with gap control implicitly embedded in the traffic flow model) are implemented to CAVs, while LCR, VSLR, and RM are issued to RHVs.

5.2. Simulation results

5.2.1. Scenario 1

Case 1, baseline: In the beginning, since there is no merging flow from the on-ramp during 0-5 min, and the lane-specific mainstream flow 2700 P.C.U./hour/lane is lower than the corresponding nominal capacity 3301 P.C.U./hour/lane under 66.7% penetration rate of CAVs. Therefore, a free-flow traffic state is expected during this period even without any control strategy, see Figure 10. From 5 min to 25 min, an on-ramp flow of 600 P.C.U./hour/lane intends to merge into lane 1 by the end of test section, see Figure 9. At the merging point connecting lane 1, **the total demand is close to the nominal capacity**, see Table 5.1. **The weaving area is operating at the level of service F that breakdown in flow would happen.** A bottleneck is thus formed. Since no control is implemented in this case, it is assumed in this case that the CAVs follow the non-optimized LCC, i.e., traversing along the closed lane until they get closed to the bottleneck. When the vehicles merge into the freeway from the ramp, they would travel at a low speed and seek for a chance to execute MLC maneuvers. When an MLC maneuver is executed by such a vehicle with low speed, the speed of lane 1 is also brought down. As a consequence, the density of lane 1 increases thus causing congestion in lane 1 and triggers lane changing maneuvers from lane 1 to lane 2 for seeking high speed. As revealed in the literature, the speed reduction and the lane changing maneuvers further induce the **capacity drop** effect, i.e., the capacity at the bottleneck is lower than the theoretical value especially when the lane-changing flow ratio is large and traffic oscillations (Chen et al., 2012; Pan et al., 2016; Zhang and Ioannou, 2017).

Speed oscillation[7] caused by the merging flow and lane changing maneuvers can be observed around the weaving area of both lanes as indicated in Figure 10. Under saturated traffic condition, when a low-speed vehicle changes to the

[7]With the introduction of CAVs, the frequency of speed oscillation becomes lower, even disappear if the traffic is composed of pure CAVs because autonomous driving can achieve speed coordination, a faster lane changing execution and require a smaller minimum safety gap. Meanwhile, the congestion is alleviated quicker than the counterpart under the pure RHVs scenario. Interested readers are referred to Pan et al. (2021) for detailed discussion.

target lane, the following vehicle traveling on the target lane must slow down to adapt to the speed of the lane-changing vehicle to guarantee traffic safety. But the speed would quickly recover when the merging vehicle increases its speed after the completion of the lane changing maneuver. The repeat of this phenomenon induces the speed oscillation. On the other hand, there is the so-called "friction effect" from the empirical observation that drivers are fear of moving fast when an incident or slowly moving vehicles exist in adjacent lanes. Due to the capacity drop, speed oscillation and lane changing behavior, the queue on lane 1 is formed and quickly spills longitudinally (on the same lane) and laterally (to the parallel lane) as shown in Figure 10. Under the baseline case, the cost incurred per vehicle is 2.57 USD/P.C.U..

Case 2, integrated optimal control strategies: To improve the freeway performance, we apply the proposed integrated optimal control strategies. Figure 11 shows that the integrated optimal control strategies can significantly improve the freeway performance. The congestion is alleviated via 1) reducing the traffic arriving at the bottleneck, i.e., via optimal VSLR/VSLC in the freeway, see Figure 12; 2) adjusting the longitudinal distribution of the lane-changing demand along the horizontal distance of the freeway segment, i.e., via optimal LCC/LCR; 3) coordinating speed and improving efficiency for the MLC maneuvers during merging and lane changing progress; and 4) ramp metering control of the on-ramp traffic.

As reported in the literature, the average speed of RHVs is roughly equal to the average speed of CAVs in a traffic stream (Bekiaris-Liberis et al., 2016; Fountoulakis et al., 2017) while the traffic is assumed to be homogeneous within one cell-lane. Therefore, the VSL controls issued to both types of vehicles are the same. The VSL is activated to reduce the flow rate arriving at the weaving area to prevent it from being an activate bottleneck, i.e., the 40 miles/hour VSL can reduce the capacity from 3301 P.C.U./hour/lane to 2830 P.C.U./hour/lane for the traffic mixed with 66.7% penetration rate of CAVs. Spatial and temporal variation of the VSL is depicted in the left hand side of Figure 12 while the corresponding fundamental diagram under speed limit control is shown in the right figure for the case of 66.7% penetration rate of CAVs. The price for this is the delay induced by slowing down the traffic at the upstream segments, which is included in the objective function defined in Equation (2). After the rush hour, the VSL gradually recovers back to the free flow speed. Lane speeds can be inhomogeneous under VSL control but without LCC (Zhang and Ioannou, 2017), which would trigger MLC and/or DLC maneuvers.

The lane-changing control is deployed to alleviate the bottleneck impact. The CAVs receive the LCC from the the VACS and execute the lane changing maneuvers far away from the bottleneck location. RHVs receive LCR from the locations where the VMS gantries are installed. The LCR triggers the MLC maneuvers for RHVs, while the execution of the these MLC follows an exponential-like function along the longitudinal dimension as discussed in (9). In contrast to Case 1, most of the lane changing maneuvers are executed at locations far away from the bottleneck location. As congestion has not yet spilled back to these locations (where lane changing maneuvers are executed), the traffic speed is still high. Therefore, the capacity drop at the incident location would be alleviated. This result is consistent with that reported in the literature (see, e.g. Zhang and Ioannou (2017); Tian et al. (2019) and the references therein). Tian et al. (2019) showed that up to 8% total travel times *of connected vehicles with lane selection assistance application* can be reduced by LCC.

Compared with Case 1, the traffic flow under Case 2 has a much lower level of congestion in terms of both the congested longitudinal distance and the time span that the congestion lasts. This is because most vehicles change their lane far away from the bottleneck spot because of the LCR/LCC implemented. Traffic condition of lane 2 is significantly improved. These observations prove that LCRs/LCCs can reduce the number of vehicles queuing at the locations at which they are forced to change lanes at low speed. Moreover, the VSLCs/VSLRs can reduce the number of vehicles arriving at the bottleneck that the possible number of vehicles seeking a chance to execute MLC maneuvers is further reduced. Due to the the speed harmonization induced by the VSLC and the minimum safety gap control (the minimum safety gap varies with respect to the real-time traffic condition and the vehicle type), the traffic can merge from the ramp in a smoother manner. Therefore, the capacity drop caused by slow lane changing traffic can be further alleviated. From a mathematical point of view, LCR/LCC make the system continuous and easier for the VSL controller (with the minimum safety gap control) to stabilize. For traffic safety, the integration of VSL and LCC controls can dramatically decrease the average number of stops, therefore drastically reducing the instances of stop-and-go traffic, smoothing the traffic flow, and damping the shockwaves. With the control scheme outlined in Case 2, the objective function significantly reduced by around 40% compared with Case 1. The average travel cost of each individual vehicle reduces from 2.57 USD/P.C.U. (Case 1) to 1.54 USD/P.C.U. (Case 2).

The oscillation of traffic speed on lane 1 around the bottleneck location is subject to turbulence caused by the ramp

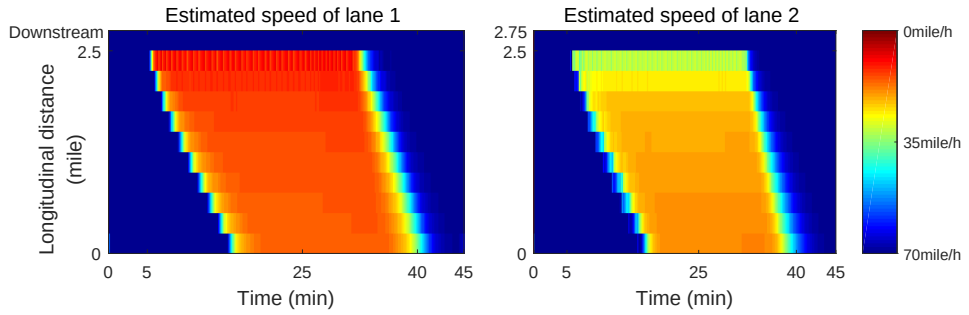


Figure 10: Spatial-temporal speed heat maps for Scenario 1 without optimal control

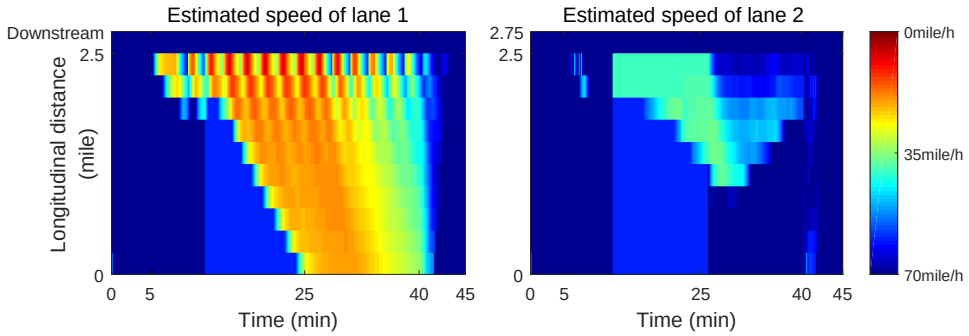


Figure 11: Spatial-temporal speed heat maps for Scenario 1 with integrated optimal control

metering. Since RHVs cannot make a cooperative decision with CAVs, the conventional signal-based ramp metering control is adopted. The switch of the green light and the red light will cause discontinuous merging flow into the freeway corridor. If the traffic condition is light, no turbulence will be observed. If the freeway corridor is already under heavy traffic conditions, this discontinuous merging flow will disrupt the mainstream traffic via lane changing maneuvers with a very small gap and under low speed. This is a typical traffic phenomenon observed in merging traffic around the on-ramp under heavy traffic conditions. There will be no interruption to the mainstream traffic from the on-ramp when there is no merging flow from the on-ramp (i.e., red signal). Therefore, the turbulence can be observed due to this discontinuous merging flow caused by the ramp control traffic signal under heavy traffic conditions. It is found in the literature that ramp control could not perform well under heavy traffic conditions with low traffic speed (Papageorgiou and Kotsialos, 2002; Hegyi et al., 2005a,b). To verify the turbulence is induced by the ramp control, we take out the ramp metering control. The simulation result is depicted in Figure 13. The oscillation of traffic speed on lane 1 around the bottleneck location no longer persists. However, compared with Case 2, traffic speeds decrease on both lanes. The average travel cost of each individual vehicle is about 1.60 USD/P.C.U., a little bit higher than that of Case 2. Compared with Case 1, the improvement is about 37.7%. This improvement is due to the integration of speed harmonization (with minimum safety gap control) and LCC as previously explained. We speculate that the ramp metering would be less effective when the penetration rate of CAVs is high, which will be further analyzed in the sensitivity analysis.

5.2.2. Sensitivity analysis

In this section, sensitivity analysis is firstly conducted to demonstrate the performance of the integrated optimal control against the penetration rate and the demand level. Then, the effect of ramp metering is studied against penetration rate. Finally, the response time of CAVs is considered as another sensitive factor that affect the performance of integrated optimal control.

The effect of integrated control against penetration rate: The results regarding the system performance with respect to different demand levels and penetration rates of CAVs are presented in Table 6. We termed Scenario 2 and

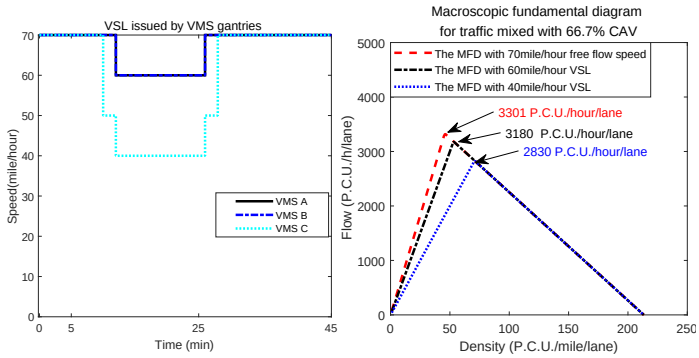


Figure 12: Variable speed limit control and the fundamental diagram

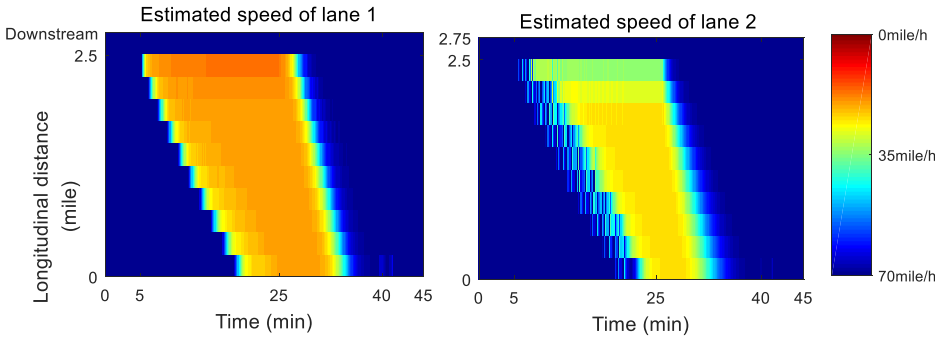


Figure 13: Spatial-temporal speed heat maps for Scenario 1 with integrated optimal control (No ramp metering)

Scenario 3 as medium demand and low demand cases, respectively. However, the penetration rates of CAVs for these two cases are also much lower than that of Scenario 1. Since the nominal capacity is positively related to the penetration rate of CAVs, and thus the definitions of heavy traffic and light traffic conditions are differentiated with respect to different penetration rates of the CAVs. Compared with the nominal capacity under the corresponding penetration rate, the total demand is close to the nominal capacity at the merging point connecting lane 1, see Table 5.**Without optimal control, the weaving area is operating at the level of service F that breakdown in flow would happen for all three scenarios**. From Table 6, we can observe that the improvement achieved by the integrated optimal control would be more significant if the penetration rate is higher, i.e., the improvement rate increases from 17.7% to 29.7% and then 40% with the increment of penetration rate of CAVs from 0% to 33.33% and 66.7%.

Next, the average travel cost against the penetration rate of CAVs is depicted in Figure 14 for all three scenarios. Under the same inflow demand, the individual travel cost is monotone decreasing with respect to the increment of penetration rate of the CAVs with or without the integrated optimal control. The CAVs can significantly improve traffic efficiency even under dense traffic conditions. The minimum individual travel cost is achieved when all vehicles are CAVs. Secondly, the proposed integrated optimal traffic control can effectively decrease the individual travel cost for all scenarios. The higher travel demand (and also the higher penetration rate of CAVs) is, the larger improvement can be observed.

The effect of integrated control against response time of CAVs: In numerical example, the response time of CAVs Δ_{CA} is 0.35seconds. Note that the response time of CAVs is assumed to vary from 0.25 sec to 1.5 sec in Levin and Boyles (2016a). Figure 15 demonstrates the system performance and the effect of optimal control against three different values of response time, i.e., 0.35 seconds, 0.8 seconds and 1.2 seconds. As it can be inferred, the individual travel cost will significantly increase with the increment of response time of CAVs for all scenarios.

The effect of ramp metering against penetration rate of CAVs: In the previous section, we noticed that the ramp metering control under high travel demand can cause oscillation of traffic speed on lane 1 around the bottleneck location. Also we speculated that the ramp metering would be less effective when the penetration rate of CAVs is

Table 6: System performance against demand level and penetration rate of CAVs

Mainline inflow (P.C.U./hour /lane)	On-ramp inflow (P.C.U./ hour)	Penetration rate	Case	Travel cost on mainstream (USD)	Extra-queue and ramp cost (USD)	Fuel and emission (USD)	Electricity (USD)	Penalty (USD)	Cumulative Flow (P.C.U.)	Average individual cost (USD/P. C.U.)
2700 High demand	600	66.7% Scenario 1	Case 1	6002	2384	219	222	922	3800	2.57
			Case 2	4573	855	219	210	0		1.54 (-40.0 %)
1800 Medium demand	399	33.3% Scenario 2	Case 1	4411	369	290	74	0	2553	2.03
			Case 2	2655	591	336	74	0		1.44 (- 29.7%)
1350 Low demand	300	0% Scenario 3	Case 1	3133	154	351	0	0	1900	1.92
			Case 2	2210	426	365	0	0		1.58 (-17.7%)

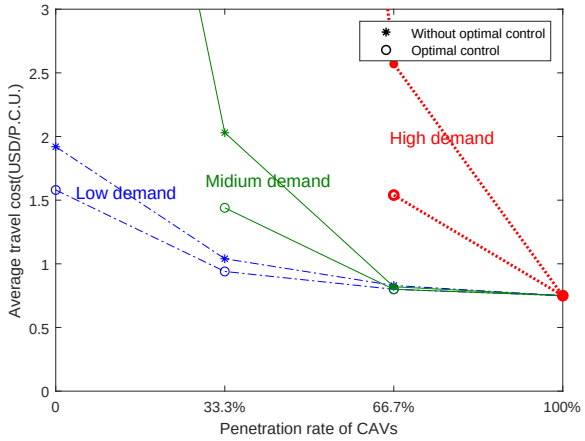


Figure 14: Average travel cost of each vehicle against penetration rate and level of demand

high. Here, we would like to investigate the effect of ramp metering control with respect to different penetration rates of CAVs and demand levels. In the first test, the bottleneck travel demand levels are plotted against the nominal fundamental diagrams with respect to the corresponding penetration rates of CAVs, respectively. Note that the weaving area would be congested if no control strategy is implemented for all these four scenarios, see Figure 16(a). The improvement rate in Figure 16(b) is thus calculated

$$\text{Improvement rate} = \frac{\text{Individual cost without ramp metering} - \text{Individual cost with integrated optimal control}}{\text{Individual cost without ramp metering}} \times 100\%$$

From Figure 16(b), the improvement rate achieved by the ramp metering control is monotone decreasing with respect to the penetration rate of CAVs and the demand level. For another comparison, we evaluate this improvement rate with respect to the penetration rate of CAVs under the same demand level, i.e., mainline 1800 P.C.U./hour/lane and 399 P.C.U./hour from the on-ramp. Note that, under the same demand level, the bottleneck congestion level can be varying with respect to the penetration rate of CAVs, see Figure 16(a). This demand pattern can cause congestion for penetration rates of 11.1% and 33.3% while free-flowing for penetration rates of 50% and 66.7%. Figure 17 depicts the effect of ramp metering control with respect to the penetration rate of CAVs and congestion level under the same demand pattern. It is found that the ramp metering control can be effective when the RHVs are dominating the traffic (and under over-saturated traffic condition). When the penetration rate of CAVs increased to a sufficient level, the effect of ramp metering is negligible under this demand pattern.

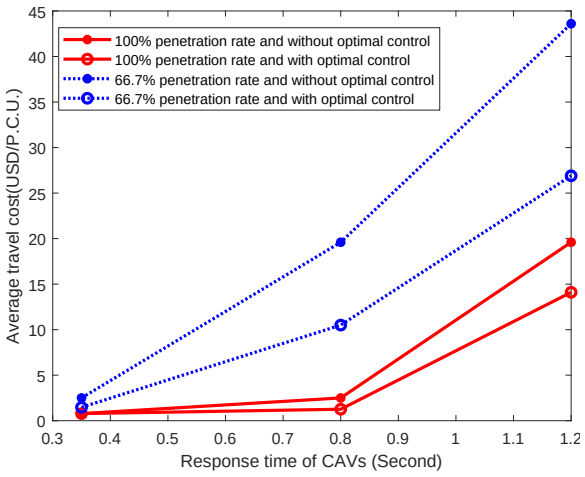
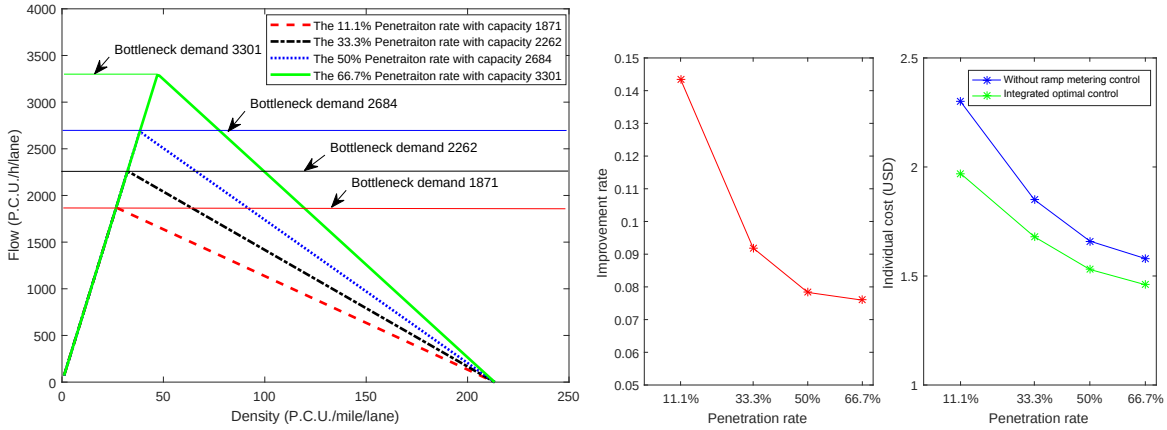


Figure 15: Average travel cost of each vehicle against response time of CAVs



(a) Bottleneck demand V.S. nominal fundamental diagram with respect to the penetration rate of CAVs

(b) Effect of ramp metering control

Figure 16: Effect of ramp metering control with respect to the penetration rate of CAVs and congestion level

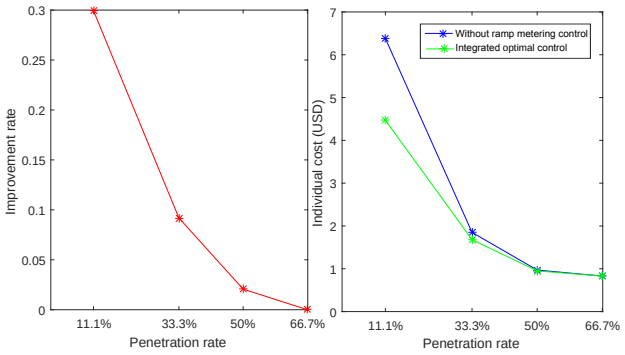


Figure 17: Effect of ramp metering control with respect to the penetration rate of CAVs under the same demand pattern

6. Conclusions

This paper proposed an integrated optimal control framework for improving the efficiency of freeway traffic mixed with CAVs and RHVs. Control measures such as the ramp metering, lane-changing and variable speed limit control (with the minimum safety gap implicitly included) are integrated to coordinate the movements of CAVs and RHVs. The controls are actuated through the VACS automatically for the CAVs. In contrast, RHVs that are not connected to the VACS are expected to make decisions in response to the control strategies and incident alarm disseminated through the en-route VMSs. The multiclass multilane CTM traffic flow model explicitly considers the variations of fundamental diagram in response to various penetration rates of CAVs and the implemented speed limit control. The compliance of drivers to the LCC is also captured by the underlying traffic flow model. Different safety gap acceptance criteria were proposed to model various lane-changing maneuvers for different vehicle types with different lane-changing intentions to guarantee traffic safety and lane-changing priorities. For practical reasons, restrictions on the control variables such as the speed limit fluctuation and lane changing frequency in both spatial and temporal dimensions were introduced to avoid excessive nuisance to drivers while maintaining traffic flow stability.

Because of the complexity of the multiclass multilane traffic flow model and the set of constraints, the optimal control problem is with non-differentiable complicated functional structure, which would induce many local optima. To tackle such challenges, a reinforcement learning based gradient-free CEM algorithm was developed for searching the optimal control. Control parameterization technique was applied to approximate the control function to yield a low dimensional parameter optimization approximation of the original high dimensional optimal control problem. Instead of searching for the value function by iteration, this gradient-free solution algorithm first parameterizes the control (or policy) and then searches for the optimal parameters that would lead to maximal returns (i.e., minimizing the objective function).

From the numerical simulation, it was found that the LCCs were able to reduce the number of vehicles queuing at the bottleneck and decrease the traffic density on the freeway. From a mathematical point of view, LCCs can make the freeway traffic flow continuous and easier for the VSL controller to stabilize. In terms of road traffic safety, the integration of VSL and LCC can drastically reduce the instances of the stop-and-go traffic and suppress the impact of the shockwaves on the freeway sections concerned. Finally, the effects of the penetration rate of CAVs and the level of congestion were investigated by sensitivity analysis in this paper. The advantage of CAVs in reducing average travel cost was highlighted. The dynamic proportion of CAVs would significantly affect the throughput and traffic speed of the freeway segment concerned. It was also found that when the penetration rate of CAVs is high enough, ramp metering does not necessary exist.

A future research direction is to extend the proposed optimal control framework to consider robustness against the inherent uncertainties in the penetration rate of CAVs, variations of the freeway supply functions such as capacity, stochastic compliance rate of drivers and dynamic delays prevailing in the control loop.

Acknowledgments

Financial support from the National Key R&D Program of China (No. 2018YFB1600500) and the Consulting project of Chinese Academy of Engineering (Nos. 2019-XZ-4 & 2019-XZ-55) is gratefully acknowledged.

Appendix: Headway distribution and fundamental diagram

To make the paper self-contained, we present a discussion on the penetration of CAVs on the headway distribution and freeway fundamental diagram, which is a recapitulation of the context in Pan et al. (2021), in this appendix.

Headway, defined as the time/space between the same positions of two consecutive vehicles, is an important measure of traffic flow characteristics, and thus it is essential for studying traffic flow. Because CAVs have significantly different operating characteristics compared to RHVs, it is deemed that CAVs can significantly reduce the headway between vehicles and hence increase the roadway capacity (Levin and Boyles, 2016a,b). A recent experiment at the California Partners for Advanced Transportation Technology (PATH) showed that CAVs in platoons can maintain a time headway as small as 0.6 s, compared to 1.5 s for RHVs (Chen et al., 2017). In view of this, it is necessary to

define a headway distribution law to model the potential capacity enhancement with the introduction of CAVs and to enable minimum headway control for mixed traffic.

In line with Levin and Boyles (2016a,b); Chen et al. (2017), the effect of CAVs on roadway capacity was investigated by considering a single-lane freeway segment. Suppose a platoon of vehicles is traveling along a freeway section in the same lane and that the traffic flow is stable without interruption from on-ramp/off-ramp or traffic incidents. Based on the rear-end collision avoidance principle (Jepsen, 1998; Levin and Boyles, 2016a,b), for vehicles traveling at speed $v(k)$ (miles/hour) the space headway (from the head of the leading vehicle to the head of the following vehicle) criterion $H_d(k)$ (mile) ahead of a specific vehicle belongs to vehicle class d is defined as below:

$$H_d(k) \geq v(k)\,\Delta T_d + l + C \quad \text{for } d = \{\text{CA}, \text{RH}\} \tag{30}$$

where ΔT_{CA} and ΔT_{RH} (hour) denote the response times of CAVs and RHVs, respectively, l (mile) denotes the vehicle length of the leading vehicle, and C is the safety gap (Jepsen, 1998) or minimum safe constant gap when all the vehicles are at a standstill (Hidas, 2005). Compared with RHVs, CAVs can tolerate a much smaller space headway because of the smaller response time. Supposing that the traffic on this unit length freeway segment (otherwise, multiply both sides of Equation (31) by the segment length) with a proportion, $P(k)$, of CAVs and RHVs, $1 - P(k)$, at time k, the relationship between space headway and traffic density is described as

$$\rho(k)\left(P(k)\,H_{\text{CA}}(k) + (1 - P(k))\,H_{\text{RH}}(k)\right) = 1 \tag{31}$$

where $\rho(k)$ is the traffic density, $H_{\text{CA}}(k)$ and $H_{\text{RH}}(k)$ denote the space headway of CAVs and RHVs, respectively.

Based on Equation (30), we have the following relations:

$$\begin{aligned}
&\rho(k)\,P(k)\,H_{\text{CA}}(k) \geq \rho(k)\,P(k)\,(v(k)\,\Delta T_{\text{CA}} + l + C)\\
&\rho(k)\,(1 - P(k))\,H_{\text{RH}}(k) \geq \rho(k)\,(1 - P(k))\,(v(k)\,\Delta T_{\text{RH}} + l + C)
\end{aligned}$$

Summing the left and right sides, respectively, we obtain

$$\rho(k)\left(P(k)\,H_{\text{CA}}(k) + (1 - P(k))\,H_{\text{RH}}(k)\right) \geq \rho(k)\,P(k)\,(v(k)\,\Delta T_{\text{CA}} + l + C) + \rho(k)\,(1 - P(k))\,(v(k)\,\Delta T_{\text{RH}} + l + C)$$

This in conjunction with Equation (31) provides

$$1 \geq \rho(k)\,P(k)\,(v(k)\,\Delta T_{\text{CA}} + l + C) + \rho(k)\,(1 - P(k))\,(v(k)\,\Delta T_{\text{RH}} + l + C)$$

Therefore, the maximum speed $\tilde{v}_s(k)$ that can be evaluated by specific traffic density and proportion of CAVs is evaluated as follows:

$$\tilde{v}_s(k) = \frac{(1 - l\rho(k) - C\rho(k))}{\rho(k)} \frac{1}{(P(k)\,\Delta T_{\text{CA}} + (1 - P(k))\,\Delta T_{\text{RH}})} \tag{32}$$

without considering the speed limit or vehicles' mechanical capability. As reported in the literature, the average speed of RHVs is roughly equal to the average speed of CAVs in a traffic stream (Bekiaris-Liberis et al., 2016; Fountoulakis et al., 2017). To this end, it is assumed that both CAVs and RHVs will spontaneously follow this maximum speed as it guarantees the minimum space gap for avoiding collisions in accordance with different levels of congestion specified by traffic density. Therefore, in this section, the variable $\tilde{v}_s(k)$ is named the **spontaneous speed limit** (SSL). However, when the traffic density tends to zero (i.e., no vehicle is traveling on the freeway), the SSL approaches an infinitely large value. To remedy this, one can simply set an upper bound to this SSL, called a vehicle's **maximum mechanical speed**. For example, 170 miles/hour is a typical upper bound of the speedometer of private vehicles. Generally speaking, the speed is actually restricted by the posted **permanent compulsory** (upper bound) **speed limit** for freeway traffic management purpose. In the United States under normal conditions, the posted permanent compulsory freeway speed limit ranges from 90 miles/hour in rural areas to 40 miles/hour in urban areas. Additionally, a temporary VSL can be issued as a control strategy for traffic incident management or congestion resolution if needed. As it can be expected, **maximum mechanical speed > permanent compulsory speed limit > temporary VSL**; therefore, the **implemented speed limit (ISL)** is generally given as either the permanent compulsory speed limit or the temporary VSL.

In line with the VSL control literature, e.g., Hegyi et al. (2005a,b), the turning point of ISL and SSL is located at the critical density $\rho^{cr}(k)$ (noting that the critical density is also affected by the implemented speed limited control

itself). From the above analysis, the traffic flow speed $v(k)$ is finally defined as a function of ISL (the lower of the posted compulsory speed limit and the VSL), traffic density, and the proportion of CAVs.

$$v(k)=\begin{cases}\tilde{v}_l(k), & \text{if } \rho(k)\le\rho^{cr}(k)\\ \tilde{v}_s(k)=\frac{(1-l\cdot\rho(k)-C\cdot\rho(k))}{\rho(k)}\frac{1}{\left(P(k)\Delta T_{\mathrm{CA}}+(1-P(k))\Delta T_{\mathrm{RH}}\right)}, & \text{otherwise}\end{cases} \tag{33}$$

where $\tilde{v}_l(k)$ denotes the ISL (which can be regarded as a decision variable of the dynamic optimization problem or the permanent compulsory speed limit posted by freeway management center) and $\rho^{cr}(k)$ is a function of the penetration rate of CAVs and ISL as follows:

$$\rho^{cr}(k)=\frac{1}{\tilde{v}_l(k)\left(P(k)\ \Delta T_{\mathrm{CA}}\ +(1-P(k))\ \Delta T_{\mathrm{RH}}\right)+l+C} \tag{34}$$

In traffic flow theory, the flow is the product of speed, $v(k)$, and density, $\rho(k)$, whereas the capacity is the maximum traffic flow (rate) observed at the critical density $\rho^{cr}(k)$. By $Q_m(k)=\rho^{cr}(k)\,\tilde{v}_l(k)$, the roadway capacity can be defined below as a function of the penetration rate of CAVs and ISL:

$$Q_m(k)=\frac{\tilde{v}_l(k)}{\tilde{v}_l(k)\left(P(k)\,\Delta T_{\mathrm{CA}}+(1-P(k))\,\Delta T_{\mathrm{RH}}\right)+l+C} \tag{35}$$

For the congested part, according to Equation (33), the flow can be evaluated as

$$\rho(k)\,\tilde{v}_s(k)=\rho(k)\,\frac{(1-l\cdot\rho(k)-C\cdot\rho(k))}{\rho(k)}\cdot\frac{1}{\left(P(k)\,\Delta T_{\mathrm{CA}}+(1-P(k))\,\Delta T_{\mathrm{RH}}\right)}=\frac{1-(l+C)\rho(k)}{\left(P(k)\,\Delta T_{\mathrm{CA}}+(1-P(k))\,\Delta T_{\mathrm{RH}}\right)}$$

Assuming a linear relationship in the congested part, we have

$$w_c(k)\left(\rho_J-\rho(k)\right)=w_c(k)\left(\frac{1}{l+C}-\rho(k)\right)=\frac{1-(l+C)\rho(k)}{\left(P(k)\,\Delta T_{\mathrm{CA}}+(1-P(k))\,\Delta T_{\mathrm{RH}}\right)}=\frac{l+C}{\left(P(k)\,\Delta T_{\mathrm{CA}}+(1-P(k))\,\Delta T_{\mathrm{RH}}\right)}\left(\frac{1}{l+C}-\rho(k)\right)$$

Therefore the back-wave speed is evaluated by:

$$w_c(k)=\frac{l+C}{\left(P(k)\,\Delta T_{\mathrm{CA}}+(1-P(k))\,\Delta T_{\mathrm{RH}}\right)} \tag{36}$$

while the jam density is determined by the average vehicle length and the minimum safe constant gap $\rho_J=\frac{1}{l+C}$. To ensure safety, the minimum space headway criteria $\tilde{H}_d(k)$ for $d=CA\ and\ RH$ are evaluated respectively as

$$\begin{aligned}\tilde{H}_{\mathrm{CA}}(k)&=v(k)\,\Delta T_{\mathrm{CA}}+l+C\\ \tilde{H}_{\mathrm{RH}}(k)&=v(k)\,\Delta T_{\mathrm{RH}}+l+C\end{aligned} \tag{37}$$

The total space $\tilde{O}(k)$ reserved for the minimum-space headway of all vehicles traveling on the freeway segment is calculated according to the proportions of the two vehicle classes.

$$\tilde{O}(k)=\rho(k)\,P(k)\,\tilde{H}_{\mathrm{CA}}(k)+\rho(k)\,(1-P(k))\,\tilde{H}_{\mathrm{RH}}(k) \tag{38}$$

When $\rho(k)\ge\rho^{cr}(k)$, the speed is determined by SSL $\tilde{v}_s(k)$ according to Equation (33), except for the total space reserved for the minimum-space headway $\tilde{O}(k)$, no other empty space is available. That is, the relationship between the traffic speed and density attains a critical value. The flow speed cannot be increased so as to maintain safety. However, when $\rho(k)<\rho^{cr}(k)$, ISL $\tilde{v}_l(k)$ can be lower than SSL $\tilde{v}_s(k)$, thus allowing extra space.

$$\tilde{O}(k)\begin{cases}<1 & \text{if } \rho(k)<\rho^{cr}(k)\\ =1 & \text{if } \rho(k)\ge\rho^{cr}(k)\end{cases}$$

That is to say, if there is no ISL $\tilde{v}_l(k)$, vehicles can travel faster under light traffic conditions. The free space $1-\tilde{O}(k)$ not occupied by vehicles can be randomly distributed among them. For simplification, we assume the safe space

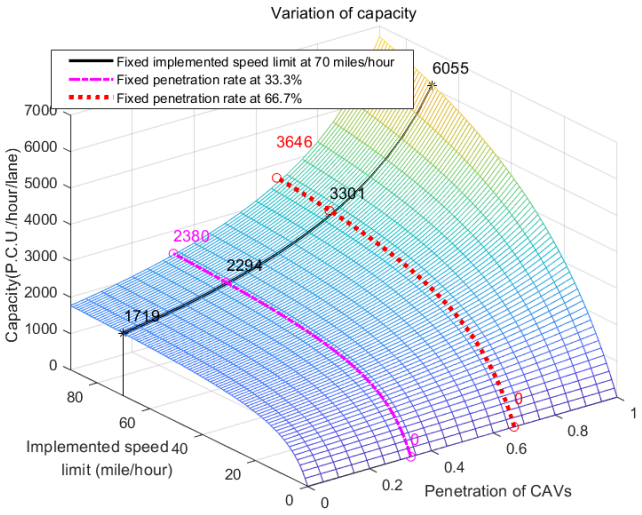


Figure 18: Capacity as a function of the penetration rate of CAVs and implemented speed limit

headway criteria are magnified by $\frac{1}{\tilde{O}(k)}$, i.e.,

$$\begin{array}{l} H_{\mathrm{CA}}(k) = \tilde{H}_{\mathrm{CA}}(k) / \tilde{O}(k) \\ H_{\mathrm{RH}}(k) = \tilde{H}_{\mathrm{RH}}(k) / \tilde{O}(k) \end{array} \quad \text{if } \rho(k) < \rho^{cr}(k)$$

To sum up, the (average) space headway distributions are thus calculated:

$$H_{\mathrm{CA}}(k) = \begin{cases} v(k)\Delta T_{\mathrm{CA}} + l + C, & \text{if } \rho(k) \geq \rho^{cr}(k) \\ \frac{v(k)\Delta T_{\mathrm{CA}} + l + C}{\rho(k)\cdot(l+C) + \rho(k)\cdot v(k)\cdot(P(k)\Delta T_{\mathrm{CA}} + (1-P(k))\Delta T_{\mathrm{RH}})}, & \text{otherwise} \end{cases}$$

$$H_{\mathrm{RH}}(k) = \begin{cases} v(k)\Delta T_{\mathrm{RH}} + l + C & \text{if } \rho(k) \geq \rho^{cr}(k) \\ \frac{v(k)\Delta T_{\mathrm{RH}} + l + C}{\rho(k)\cdot(l+C) + \rho(k)\cdot v(k)\cdot(P(k)\Delta T_{\mathrm{CA}} + (1-P(k))\Delta T_{\mathrm{RH}})}, & \text{otherwise} \end{cases}$$

The brake response time of a human driver includes mental processing, muscle movement, and brake engagement time. On average, the mental processing takes about 1.3 s for unexpected occasions, the average muscle movement takes 0.2 s, and the brake engagement time takes 0.35 s under emergency conditions. Therefore, the braking response time of RHVs to an unexpected occasion is 1.85 s. However, CAVs do not have mental reactions nor muscle movement; therefore, the response time of CAVs is considered to be 0.35 s for the brake engagement process. This value is consistent with the 0.6 sec time headway between successive CAVs (Chen et al., 2017).

Figure 18 demonstrates the impact of the penetration rate of CAVs and the compulsory speed limit on road capacity. In this example, all vehicles are considered passenger car equivalent (P.C.E.) or passenger car unit (P.C.U.) vehicles 20 ft in length, and the safe constant gap is 6.5 ft (when the related vehicles are at a standstill). As stated above, the response times of RHVs and CAVs are set to be 1.85 sec and 0.35 s, respectively. As shown in Figure 18, the capacity monotonically increased with the increasing penetration rate of CAVs and the ISL. The solid black line quantifies the variation of capacity with respect to the penetration rate of CAVs varying from 0% to 100% by fixing the compulsory speed limit at 70 miles/hour. As demonstrated in this figure, the capacity varies significantly from 1719 P.C.U./hour/lane (with 100% RHVs) to 6055 P.C.U./hour/lane (with 100% CAVs). On the other hand, by conventional macroscopic traffic flow theory, the increase in free-flow speed (or the compulsory speed limit in our case) would introduce an increase in capacity. The red dotted line presents an example for this using a fixed 66.7% penetration rate of CAVs, while the pink dot-dash line presents the 33.3% one.

References

Ahn, S., Cassidy, M., 2007. Freeway traffic oscillations and vehicle lane-change maneuvers. In: Allsop, R.E., Bell, M.G.H., Heydecker, B. (Eds.), Proceedings of the 17th ISTTT, 691-710.

Bekiaris-Liberis, N., Roncoli, C., and Papageorgiou, M., 2016. Highway traffic state estimation with mixed connected and conventional vehicles. IEEE Trans. Intell. Transp. Syst., 17(12), 3484-3497.

Bertsekas, D., 2005. Dynamic programming and optimal control. Athena Scientific, Belmont, MA.

Busoniu, L., Babuska, R., De Schutter, B., Ernst, D., 2010. Reinforcement learning and dynamic programming using function approximators. CRC Press, Boca Raton, FL.

Carlson, R., Papamichail, I., Papageorgiou, M., Messmer, A., 2010. Optimal motorway traffic flow control involving variable speed limits and ramp metering. Transp. Sci., 44 (2), 238-253.

Carlson, R., Papamichail, I., Papageorgiou, M., 2011. Local feedback-based mainstream traffic flow control on motorways using variable speed limits. IEEE Trans. Intell. Transp. Syst., 12(4), 1261-1276.

Chen, D., Laval, J., Zheng, Z., Ahn, S. 2012. A behavioral car-following model that captures traffic oscillations. Transport. Res. B, 46(6), 744-761.

Chen D., Ahn, S., Chitturi, M., and Noyce, D., 2017. Towards vehicle automation: Roadway capacity formulation for traffic mixed with regular and automated vehicles. Transport. Res. B, 100, 196-221.

Diakaki, C., Papageorgiou, M., Papamichail, I., Nikolos, I., 2015. Overview and analysis of Vehicle Automation and Communication Systems from a motorway traffic management perspective. Transport. Res. A, 75, 147-165.

Fountoulakis, M., Bekiaris-Liberis, N., Roncoli, C., Papamichail, I., and Papageorgiou, M., 2017. Highway traffic state estimation with mixed connected and conventional vehicles: Microscopic simulation-based testing. Transport. Res. C, 78, 13-33.

Gabillon, V., Ghavamzadeh, M., Scherrer, B., 2013. Approximate dynamic programming finally performs well in the game of Tetris. In Proceedings of Neural Information Processing Systems (NIPS) Dec 2013, South Lake Tahoe, United States.

Han, Y., Chen D., and Ahn, S., 2017. Variable speed limit control at fixed freeway bottlenecks using connected vehicles. Transport. Res. B, 98, 113-134.

Jia, D., and Ngoduy, D., 2016a. Enhanced cooperative car-following traffic model with the combination of V2V and V2I communication. Transport. Res. B, 90, 172-191.

Jia, D., and D. Ngoduy. 2016b. Platoon based cooperative driving model with consideration of realistic inter-vehicle communication. Transport. Res. C, 68, 245-264.

Jia, D., Ngoduy, D., and Vu, H., 2019. A multiclass microscopic model for heterogeneous platoon with vehicle-to-vehicle communication. Transportmetrica B, 7(1), 448-472.

Hegyi, A., Schutter, B., Hellendoorn, H., 2005a. Model predictive control for optimal coordination of ramp metering and variable speed limits. Transport. Res. C, 13(3), 185-209.

Hegyi, A., Schutter, B., Hellendoorn, H., 2005b. Optimal coordination of variable speed limits to suppress shock waves. IEEE Trans. Intell. Transp. Syst., 6(1), 102-112.

Hidas, P., 2005. Modelling vehicle interactions in microscopic simulation of merging and weaving. Transport. Res. C 13 (1), 37-62.

Jepsen, M. 1998. On the Speed-Flow Relationships in Road Traffic: A Model of Driver Behaviour. In Proceedings of the Third International Symposium on Highway Capacity.

Khondaker, B., Kattan, L., 2015. Variable speed limit: A microscopic analysis in a connected vehicle environment, Transp. Res. C, Emerg. Technol., 58, 146-159.

Levin, M., and Boyles, S., 2016a. A multiclass cell transmission model for shared human and autonomous vehicle roads. Transport. Res. C, 62, 103-106.

Levin, M., and Boyles, S., 2016b. A cell transmission model for dynamic lane reversal with autonomous vehicles. Transport. Res. C, 68, 126-143, 2016.

Li, Z., Liu, P., Wang, W. and Xu, C., 2014. Development of a control strategy of variable speed limits to reduce rear-end collision risks near freeway recurrent bottlenecks. IEEE Trans. Intell. Transp. Syst., 15(2), 866-877.

Li, Z., Liu, P., Xu, C. and Wang, W., 2016. Optimal mainline variable speed limit control to improve safety on large-scale freeway segments. Comput-Aided Civ. Inf., 31, 366-380, 2016.

Liu, S., Hellendoorn, H., and Schutter, B., 2017. Model predictive control for freeway networks based on multiclass traffic flow and emission models. IEEE Trans. Intell. Transp. Syst., 18(2), 306-321.

Mohajerpoor, R., Ramezani, M., 2019. Mixed flow of autonomous and human-driven vehicles: Analytical headway modeling and optimal lane management. Transp. Res. C, 109, 194-210.

Pan, T., Lam, W.H.K., Sumalee, A., Zhong, R., 2016. Modelling the impacts of mandatory and discretionary lane-changing maneuvers. Transp. Res. C, 68, 403-424.

Pan, T., Lam, W.H.K., Sumalee, A., Zhong, R., 2021. Multiclass multilane model for freeway traffic mixed with connected automated vehicles and regular human-piloted vehicles. Transportmetrica A, 17(1), 1-33.

Powell, W.B., 2011. Approximate dynamic programming: solving the curses of dimensionality, 2nd edition. Wiley, Hoboken, NJ.

Papageorgiou, M., Kosmatopoulos, E., Papamichail, I., 2008. Effects of variable speed limits on motorway traffic flow. Transport. Res. Rec., 2047, 37-48.

Malikopoulos, A., Hong, S., Park, B., Lee, J., Ryu, S., 2019. Optimal control for speed harmonization of automated vehicles, IEEE Trans. Intell. Transp. Syst., 20(7), 2405-2417.

Múller, E.R., Carlson, R.C., Kraus Jr., W., Papageorgiou, M., 2015. Microsimulation analysis of practical aspects of traffic control with variable speed limits. IEEE Trans. Intell. Transp. Syst., 16 (1), 512-523.

Soriguera, F., Martínez, I., Sala, M., Menéndez, M., 2017. Effects of low speed limits on freeway traffic flow. Transport. Res. Part C, 77, 257-274.

Gao, W., Jiang, Z., Ozbay, K., 2017. Data-driven adaptive optimal control of connected vehicles. IEEE Trans. Intell. Transp. Syst., 18(5), 1122-1133.

Gao, W., Odekunle, A., Chen, Y., Jiang, Z., 2019. Predictive cruise control of connected and autonomous vehicles via reinforcement learning. IET Control Theory & Applications, 13(17), 2849-2855.
Papageorgiou, M., Kotsialos, A., 2002. Freeway ramp metering: An overview. IEEE Trans. Intell. Transp. Syst., 3 (4), 271-281.
Roncoli, C., Papageorgiou, M., Papamichail, I., 2015a. Traffic flow optimisation in presence of vehicle automation and communication systems–Part I: a first-order multilane model for motorway traffic. Transp. Res. C, 57, 241-259.
Roncoli, C., Papageorgiou, M., Papamichail, I., 2015b. Traffic flow optimisation in presence of vehicle automation and communication systems–Part II: optimal control for multilane motorways. Transp. Res. C, 57, 260-275.
Roncoli, C., Papamichail, I., Papageorgiou, M., 2016. Hierarchical model predictive control for multilane motorways in presence of Vehicle Automation and Communication Systems. Transp. Res. C, 62, 117-132.
Rubinstein, R., and Kroese, D., 2004. The cross-entropy method: a unified approach to combinatorial optimization, monte-carlo simulation and machine learning. Springer.
Szita, I., Lorincz, A., 2006. Learning Tetris using the noisy cross-entropy method. Neural Comput., 18 (12), 2936-2941.
Tang, T., Wang, T., Chen L., and Shang H., 2017. Impacts of energy consumption and emissions on the trip cost without late arrival at the equilibrium state. Physica A, 479, 341-349.
Thiery, C., Scherrer, B., 2009. Improvements on Learning Tetris with Cross Entropy. Int. Comput. Games Assoc. J., 32 (1), 23-33.
Stern, R. E., Cui, S., Monache, M.L., Bhadani, R., Bunting, M., Churchill, M., Hamilton, N., Haulcy, R., Pohlmann, H., Wu, F., Piccoli, B., Seibold, B., Sprinkle, J., Work, D. B., 2017. Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments. Transp. Res. C, 89, 205-221.
He, Z., Zheng, L., Song, L., Zhu, N., 2017. A jam-absorption driving strategy for mitigating traffic oscillations. IEEE Trans. Intell. Transp. Syst., 18 (4), 802-813.
Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., et al. 2017. Mastering the game of go without human knowledge. Nature, 550(7676): 354.
Varaiya, P., 1993. Smart cars on smart roads: problems of control. IEEE Trans. Automat. Contr., 38 (2), 195-207.
Talebpour, A., Mahmassani, H. S., 2016. Influence of connected and autonomous vehicles on traffic flow stability and throughput, Transp. Res. C, 71, 143-163.
Tian, D., Wu, G., Hao, P., Boriboonsomsin, K., Barth, M., 2019. Connected vehicle-based lane selection assistance application. IEEE Trans. Intell. Transp. Syst., 20(7), 2630-2643.
Wang, M., Daamen, W., Hoogendoorn, S.P., van Arem, B., 2016a. Connected variable speed limits control and car-following control with vehicle-infrastructure communication to resolve stop-and-go waves. J. Intell. Transp. Syst., 20 (6), 559-572, 2016.
Wang, M., Daamen, W., Hoogendoorn, S.P., van Arem, B., 2016b. Cooperative car-following control: Distributed algorithm and impact on moving jam features. IEEE Trans. Intell. Transp. Syst., 17(5), 1459-1471.
Wang, Z., Bian, Y., Shladover, S., Wu, G., Li, S., Barth, M., 2020. A Survey on Cooperative Longitudinal Motion Control of Multiple Connected and Automated Vehicles, IEEE Intell. Transp. Syst. Mag., 12(1), 4-24.
Zheng, Z., Ahn, S., Chen, D., Laval, J., 2013. The effects of lane-changing on the immediate follower: Anticipation, relaxation, and change in driver characteristics. Transp. Res. C, 26, 367-379.
Zhong, R., Sumalee, A., Pan, T., Lam, W.H.K, 2014. Optimal and robust strategies for traffic management under demand and supply uncertainties: An overview and general theory. Transportmetrica, 10 (10), 849-877.
Zhong, R., Fu, K., Sumalee, A., Ngoduy, D., Lam, W.H.K., 2016a. A cross-entropy method and probabilistic sensitivity analysis framework for calibrating microscopic traffic models. Transp. Res. C, 63, 147-169.
Zhong, R., Yuan, F., Pan, T., Chow, A., Chen, C., Yang, Z., 2016b. Linear complementarity system approach to macroscopic freeway traffic modeling: Uniqueness and convexity. Transportmetrica A, 12 (2), 142-174.
Zhang, Y., and Ioannou, P., 2017. Combined variable speed limit and lane change control for highway traffic, IEEE Trans. Intell. Transp. Syst., 18 (7), 1812-1823.
Zhu, F., Ukkusuri, S., 2015. A linear programming formulation for autonomous intersection control within a dynamic traffic assignment and connected vehicle environment. Transp. Res. C, 55, 363-378.
Zhou, M., Qu, X., Jin, S., 2017. On the impact of cooperative autonomous vehicles in improving freeway merging: a modified intelligent driver model-based approach. IEEE Trans. Intell. Transp. Syst., 18 (6), 1422-1428,.
Zhu, F., Ukkusuri, S., 2018. Modeling the proactive driving behavior of connected vehicles: A cell-based simulation approach. Comput-Aided Civ. Inf., 33 (4), 262-281.