Towards Vehicle Automation: Roadway Capacity Formulation for Traffic Mixed with Regular and Automated Vehicles

Danjue Chen[a], Soyoung Ahn[b,1], Madhav Chitturi[b], and David A. Noyce[b]

[a] Department of Civil and Environmental Engineering, University of Massachusetts Lowell
One University Avenue, Lowell, MA 01854
[b] Department of Civil and Environmental Engineering, University of Wisconsin-Madison
2205 Engineering Hall, 1415 Engineering Drive, Madison, WI 53706

Abstract

This paper provides formulations of traffic operational capacity in mixed traffic, consisting of automated vehicles (AVs) and regular vehicles, when traffic is in equilibrium. The capacity formulations take into account (1) AV penetration rate, (2) micro/mesoscopic characteristics of regular and automated vehicles (e.g., platoon size, spacing characteristics), and (3) different lane policies to accommodate AVs such as exclusive AV and/or RV lanes and mixed-use lanes. A general formulation is developed to determine the valid domains of different lane policies and more generally, AV distributions across lanes with respect to demand, as well as optimal solutions to accommodate AVs.

Key words: automated vehicles (AV), platooning, capacity, distribution policy

[1] Corresponding Author: 2304 Engineering Hall, 1415 Engineering Drive, Madison, WI 53706; Phone: 608-265-9067; Fax: 608-262-5199 (dept.); Email: sue.ahn@wisc.edu

List of Variables:

AV: automated vehicle;
RV: regular vehicle that is at level 0 automation;
s_0: critical spacing of a RV following another RV;
β^A: spacing coefficient for the lead vehicle in an AV platoon (with a RV ahead);
γ: spacing coefficient for other AVs in the platoon (i.e., with another AV ahead);
β^R: spacing coefficient for the first RV following an AV platoon;
n: AV platoon size;
m: number of RVs between two platoons;
$\bar{s}$: mean critical spacing per cycle;
α: AV proportion in the traffic stream;
u: free-flow speed;
ε: average gain of critical spacing per AV, named "AV gain";
ε_i: AV gain on lane i;
C_0: lane capacity with only RVs;
C_i: capacity on lane i, with potential RVs and AVs;
$f(\alpha, \varepsilon)$: capacity function;
p: AV penetration rate;
q_i: flow on lane i;
α_i^*: AV proportion of lane i in the capacity state;
Q_{policy}: flow corresponding to a policy, (A, R), (M, R) or (A,M);
Q_{policy}^{max}: maximum flow (i.e., capacity) for a given p for a policy;
Q_{case}^*: maximum capacity among the possible p for a policy;
p_{cric}: critical AV penetration rate;
Q_K: total flow for a K-lane highway;
Q_K^{max}: maximum flow (i.e., capacity) for a given p for a K-lane highway.
$E_i(\alpha_i)$: AV gain for lane i as a function of AV proportion α_i.

1 Introduction

Emerging automated vehicle (AV) technologies have the potential to fundamentally change driver interactions and provide tremendous opportunities to drastically improve traffic efficiency, stability and safety. Among different types of AV technologies, vehicle platooning is particularly advantageous due to its unique high performance feature: doubled (or higher) roadway capacity and significantly improved flow stability (Milanes et al., 2014; Milanés and Shladover, 2014; Shladover et al., 2012, 2010). A few pioneering field tests conducted recently have provided the very first understanding of vehicle platooning realized through cooperative adaptive cruise control and suggested very promising improvement in roadway efficiency (Bu et al., 2010; Milanes et al., 2014; Ploeg et al., 2011). Particularly, the latest experiment at the California PATH showed that vehicles in platoons can maintain a time gap as small as 0.6 s, compared to 1.5 s for conventional non-automated vehicles, which implies a substantial increase of roadway capacity and drastic congestion mitigation (Milanés and Shladover, 2014; Shladover et al., 2012).

In traffic flow research, one important problem has received much attention: how the improvement in roadway capacity will evolve as the AV technologies mature and the penetration rate gradually increases? Understanding this problem is critical for applying the emerging technologies for traffic control and transportation planning in the era of AVs (Lin and Wang, 2013; Litman, 2015; Williams, 2013; Zhou et al., 2015). Some existing studies provide valuable insights on this issue. From the perspective of vehicle mechanics, Swaroop et al. (1994) examined the impacts of platooning policy (constant time gap vs. constant space) on traffic flow instability and evaluated lane capacity with different platoon sizes and headway settings. However, the formulation of lane capacity was overly simplified. For example, they did not consider the interaction between platoons and the distribution of AVs across different lanes. Later, using simulations, a number of studies investigated changes in microscopic driving behavior in AVs, such as reaction time and acceleration/deceleration, and platoon size, and their impacts on capacity (e.g., Jerath and Brennan, 2012; Kesting et al., 2010, 2008; Talebpour et al., 2015; Talebpour and Mahmassani, 2014; Talebpour et al., 2017; Treiber et al., 2007; van Arem et al., 2006; Zhao and Sun, 2013). For example, van Arem et al. (2006) used a microscopic simulator, MIXIC, to study the impacts of vehicle platoons on the flow instability and capacity on a freeway with a lane drop. Talebpour and Mahmassani (2016) explicitly considered traffic flow with different compositions of connected and automated vehicles and found that AVs are superior in terms of improving string stability. A large proportion of these studies found that a substantial capacity improvement can be achieved with medium (or even low) penetration rates (e.g., Jerath and Brennan, 2012; Kesting et al., 2008, 2007; Treiber and Kesting, 2013). In contrast, Shaldover et al. (2012) calibrated simulations using field experiments and found that a capacity increase is marginal until the penetration rate reaches a moderate to high level (e.g., above 50%), which is consistent with the simulation outcome of van Arem et al. (2006). Using simulations, Talebpour et al. (2017) examined the impacts of reserving one lane of a four-lane highway for AVs on traffic flow dynamics and travel time reliability. It was found that throughput can be improved significantly if the AV penetration rate is greater than 30%. However, the mechanisms of the throughput improvement are unclear because complex car-following and lane-changing dynamics were assumed in the simulations.

To date, most evaluation efforts have used simulations, and very limited theoretical research has been conducted to provide a systematic formulation. To fill this gap, this paper provides a general theoretical framework to shed light on how traffic operational capacity will change with the introduction of AVs. The operational capacity (hereafter capacity to be succinct) is defined as the maximum sustainable flow on a segment, given an AV penetration rate, when demand is sufficiently high. Particularly, this paper focuses on deriving the base macroscopic capacity for mixed traffic when traffic is in equilibrium. This is a very important and necessary step toward establishing a benchmark to meaningfully understand the effects of bottlenecks and various microscopic features. This is not trivial because the base equilibrium capacity is not a fixed number but varies (dynamically) by a number of factors, which presents complexity in understanding the capacity. Our formulations take into account (1) AV penetration rate, (2) micro/mesoscopic characteristics of regular (i.e., conventional non-automated) and automated vehicles, and (3) different lane policies to accommodate AVs such as exclusive AV and/or RV lanes and mixed-use lanes, which has received very little attention in the previous research. The micro/mesoscopic vehicle characteristics (e.g., platoon size, spacing characteristics) are expressed by a single parameter, average gain in critical spacing with AVs, thereby establishing a clear connection to the macroscopic capacity. We further develop a general formulation, inclusive of all the lane policies considered in this study, to determine how AVs should be distributed across lanes, given traffic demand, AV penetration rate, and spacing characteristics of automated and regular vehicles. The analytical formulations offer important insights into valid domains of different lane policies and more generally, AV distributions across lanes with respect to demand, as well as optimal solutions to accommodate AVs.

This paper is organized as follows. Section 2 presents the capacity formulation for a single-lane highway to reveal the impact of micro/mesoscopic vehicle spacing and platooning characteristics on roadway capacity. Section 3 presents the analysis for a two-lane highway, including the capacity formulations for three specific lane policies, followed by the general formulation. Section 4 extends the formulations to a multi-lane case. Concluding remarks are provided in Section 5.

2 Capacity Formulation for Single-lane Highway

In this section, we formulate the 'physical' lane capacity of single-lane traffic consisting of regular vehicles (RVs) and AV platoons, defined as the maximum sustainable flow for given proportions of AVs and RVs in traffic streams (independent of the AV penetration rate, p). Note that the physical lane capacity is independent of the lane policy. A more detailed distinction between the *operational* capacity and *physical* capacity is given in the following section. In this study, RVs refer to vehicles at level 0 automation (i.e., no-automation) according to the definition of NHTSA (NHTSA, 2013). Specifically, the capacity is formulated in terms of micro/mesoscopic characteristics, including the AV platoon size, required inter-vehicle spacing of AVs and RVs, and the proportion of AVs.

We assume that both AVs and RVs travel at a constant free-flow speed of u until they reach their respective *critical* spacing (corresponding to capacity), below which they enter the car-following mode. The spacing here is defined as the distance between the reference points of a leader and the immediate follower (e.g., front bumper to front bumper), as in many studies (e.g., Ahn et al., (2004), Newell (2002)). To capture the difference between AVs and RVs in spacing suggested by previous field tests (Milanés et al., 2014; Shladover et al., 2010), we differentiate four different critical spacing levels, depending on the vehicle pairing (see Fig. 2-1): (1) s_0 for a RV following another RV, (2) $\beta^A s_0$ for the lead vehicle in an AV platoon (with a RV or an AV ahead), (3) γs_0 for other AVs in the platoon, and (4) $\beta^R s_0$ for the first RV following an AV platoon, in which $\beta^A > 0, \beta^R > 0$, and $\gamma > 0$. Furthermore, we expect $\gamma < 1$ with the automation capability; $\beta^A \geq \gamma$ since a lead AV is likely to at least maintain the spacing of a non-leader AV; and $\beta^R \geq 1$ since the first RV following an AV platoon is likely to at least maintain the regular spacing. Note that these parameters, though treated as fixed and deterministic in this paper, essentially represent *average* drivers/vehicles characteristics. Since we are only concerned with the critical spacing in formulating capacity, the word "critical" is dropped hereafter.

Fig. 2-1: (a) Fundamental diagram; (b) illustration of inter-vehicle spacing characteristics

For simplification, we assume that all AVs are platooned and that AV platoons exist periodically, such that the traffic stream is periodic with each cycle consisting of one n-AV platoon and m RVs. Then, the mean critical spacing per cycle is expressed as

$$\bar{s} = \frac{\beta^A s_0 + (n-1)\gamma s_0 + \tilde{m}\left(\beta^R s_0 + (m-1)s_0\right)}{n+m}. \tag{2-1}$$

where

$$\tilde{m} = \begin{cases} 1, & if\ m \geq 1 \\ 0, & if\ m = 0 \end{cases}.$$

Note that, when $m = 0$, the traffic stream only has AVs, but it still consists of periodic platoons of finite size (n), not a single infinite platoon. We believe that this is preferred in real implementation to assure (1) effective vehicle communication – the current DSRC communication range is about 300m (Nowakowski et al., 2016), and (2) safety and efficiency, particularly for lane changes and platoon adjustment (forming or de-forming).

Eqn. (2-1) can be re-written as follows:

$$\bar{s} = s_0(1 - \alpha\varepsilon),$$

where

$$\alpha = \frac{n}{n+m},$$

$$\varepsilon = \begin{cases} 1 - \gamma - \left(\frac{\beta^A - \gamma}{n} + \frac{\beta^R - 1}{n}\right), if\ 0 \leq \alpha < 1 \\ 1 - \gamma - \frac{\beta^A - \gamma}{n}, if\ \alpha = 1 \end{cases}. \tag{2-2}$$

The α denotes the AV proportion in the traffic stream, and ε represents the average gain of critical spacing per AV– named "AV gain" hereafter. The β^R related term is dropped when $\alpha = 1$ since no RVs are present. Notice that $\varepsilon < 1$ is expected because $\gamma < 1, \beta^A \geq \gamma$, and $\beta^R \geq 1$ as previously stated. Then, the capacity, C, is derived as follows:

$$C = \frac{1}{\bar{s}/u} = \frac{u}{s_0(1-\alpha\varepsilon)} = \frac{C_0}{1-\alpha\varepsilon}, \tag{2-3}$$

where C_0 denotes the lane capacity with only RVs. We refer to this as the physical lane capacity function, $f(\alpha, \varepsilon)$: i.e.,

$$C = f(\alpha, \varepsilon) = \frac{C_0}{1-\alpha\varepsilon}. \tag{2-4}$$

Clearly, C is not constant, but depends on the AV proportion, α, and the AV gain, ε. Note that ε is determined by the spacing characteristics, γ, β^A, and β^R, the AV platoon size, n, and potentially α. Particularly, if $\beta^R > 1$, ε has different values when $0 \leq \alpha < 1$ and $\alpha = 1$; but if $\beta^R = 1$, the value of ε is consistent regardless of α; see Eqn. (2-2). Additionally, ε will vary if n is dynamic. For the formulation in Section 3 and 4, it is assumed that (i) ε is independent of α and n, and (ii) ε takes a consistent value for both $0 \leq \alpha < 1$ and $\alpha = 1$. Assumption (ii) is made to simplify the formulation but it will not affect the results. For (i), section 5 will visit the issue of correlated ε and α since n could increase with α if drivers tend to form longer platoons when they see more AVs around. For a given ε, the maximum C is achieved when $\alpha = 1$ (i.e., AVs only in the lane), which is intuitive given that $\gamma < 1, \beta^A \geq \gamma$, and $\beta^R \geq 1$. Furthermore, ε is an indicator of operational efficiency; a greater ε leads to a greater capacity. From the formulation above, we obtain the following remarks about the impacts of the platoon size and the spacing characteristics on the physical lane capacity:

R1: The capacity (C) increases with platoon size (n); see Fig. 2-2(a). This is intuitive since RVs require greater spacings than the non-leader AVs: i.e., $\gamma < 1$, and additional spacing may be required for the lead AV and the first RV following an AV platoon: i.e., $\beta^A \geq \gamma$, and $\beta^R \geq 1$. The extra spacings will diminish the average AV gain by $\frac{\beta^A - \gamma}{n} + \frac{\beta^R - 1}{n}$, shown in Eqn.(2-2), which decreases with n.

R2: The capacity decreases with γ, as expected; see Fig. 2-2(b).

R3: The capacity decreases with β^A as expected; see Fig. 2-2(c). Notice that when β_A is sufficiently large, $\varepsilon < 0$ and $C < C_0$ (e.g., the blue line in the figure). The same remark holds for β^R since it has a similar characteristics. These results suggest that overly cautious driving by the lead AVs or RVs following a platoon could potentially degrade the capacity. This is particularly likely in the presence of long platoons.

R4: If $\varepsilon > 0$, the capacity increases with the AV proportion (α) at an increasing rate; see the green line on Fig. 2-2(a) for example. This is straightforward since as α increases, the average spacing ($\bar{s}$) decreases linearly, and thus capacity increases at the inverse rate.

Fig 2-2: (a) Capacity change with various n ($\gamma = 0.5, \beta^A = \beta^R = 1$); (b) Capacity change with various γ ($n = 6, \beta^A = \beta^R = 1$); (c) Capacity change with various β^A ($n = 4, \gamma = 0.5, \beta^R = 1$).

Note that, the expected parameter ranges (i.e., $\gamma < 1$, $\beta^A \geq \gamma$, and $\beta^R \geq 1$) are based on the operational features of AVs revealed in the literature (Milanés et al., 2014; Shladover et al., 2010). However, it is possible that these ranges are violated and that we have sufficiently small β^A ($< \gamma$) and β^R(< 1), and thus $\left(\frac{\beta^A - \gamma}{n} + \frac{\beta^R - 1}{n}\right) < 0$. In that case, the opposite of R1 will hold: the physical lane capacity (C) decreases with platoon size (n). Fortunately, this is well captured in ε because our formulation uses the general values of γ, β^A, and β^R as long as they are physically meaningful (i.e., non-negative).

The results obtained in this section show that the impacts of vehicles' micro/mesoscopic characteristics (including spacing characteristics and platoon size) on the physical lane capacity can be captured by a single parameter, ε. Therefore, in our next analysis, we leave out the detailed vehicle characteristics and use ε to indicate the feature of AV gain for a lane.

3 Two-lane Highway

In this section, we consider the scenario of a two-lane highway. Fig. 3-1 shows a sketch of a general framework. Here we assume that the platooning parameters introduced in the single-lane scenario in section 2 can vary across lanes. Let α_i denote the AV proportion in lane i, q_i the flow, and C_i the physical lane capacity determined by Equation (2-4) (with parameters α_i and ε_i); see the definition of the parameters provided in the beginning of the paper. The objective of this section is to derive the (operational) capacity of a two-lane highway under various lane policies. Specifically, we investigate the capacity (synonymous to "discharge flow") defined as the maximum sustainable flow on a segment when demand is sufficiently high. The capacity depends on the lane policy to distribute AVs across lanes and is also constrained by the AV penetration rate, p. Thus, the sum of the physical lane capacities across lanes is the upper bound for the capacity for a facility. Underutilization of the physical capacity can occur due to p and the treatment of vehicle entrance.

Notably, it is assumed that when traffic enters a highway (e.g., via on-ramps), the proportion of vehicles that are AVs remains at p throughout the highway. (Essentially, we assume that FIFO is maintained throughout the system even when only one of the lanes is at physical lane capacity. More discussion on this is provided in Section 6.) However, p can be regarded as the proportion of AVs entering the highway to be more general. Note that the traffic composition can change at a highway entrance if special schemes are designed to segregate different vehicle types or to prioritize a certain type. For example, if AVs and RVs are separated at the entrance (with two 'equal' lanes) and have an even chance to enter the highway, then the AV proportion entering the highway will be fixed at 0.5 (assuming sufficiently high demand). In this case, the formulations in this paper will still apply to the simple case of p = 0.5. Similarly, if AVs are prioritized to have guaranteed entrance from the ramp (in low penetration), a lane that allows for AVs (not necessarily exclusive) is *always* able to take sufficient AVs from the demand to reach its physical lane capacity (given by Eqn. (2-4)). Under such prioritization scheme, the capacity of a two-lane highway would be equal to the sum of physical lane capacity regardless of the lane policy.

To simplify derivations, we assume that in the absence of AVs, the two lanes have the same physical lane capacity, C_0. Since C_0 is a constant, it does not affect the behavior of capacity with respect to more important parameters, α and ε. It is also assumed that lane 1 has a higher priority in platooning and thus the AV gain in lane 1 (left lane) is greater than lane 2; i.e., $\varepsilon_1 > \varepsilon_2$. Such higher priority is possible for various reasons, such as preference of AVs in the left lane or stricter regulation of platoon length in the shoulder lane (i.e., lane 2) to favor vehicle merging or exit. Nevertheless, later in section 4, we will relax this assumption and derive a more general formulation.

Below, we first derive the capacity functions for three different lane policies possible as the AV technology matures and the penetration rate gradually increases. Then, a more general formulation will follow to determine the feasible policies and more generally, valid domains of AV lane distributions for various levels of demand.

Fig. 3-1: Sketch of two-lane framework.

3.1 Capacity functions under different lane policies

Here we study the capacity under three different lane policies: (1) (A, R) policy, in which lane 1 only allows AV platoons and lane 2 RVs; (2) (M, R) policy, in which lane 1 allows both AV platoons and RVs (i.e., mixed traffic), and lane 2 is still RVs-only; and (3) (A, M) policy, in which lane 1 is dedicated to AV platoons and lane 2 has mixed traffic. The (A, R) policy is likely the most preferable for a smooth AV transition to avoid driver confusion and minimize safety risk. However, this policy could significantly undermine the capacity if the AV penetration rate is too low or high. The (M, R) policy represents a compromise in low AV penetration that limits AV platooning to certain lane(s) to reduce safety risks while fully realizing the capacity enhancement. When the AV penetration becomes sufficiently high, the (A, M) policy would be desired, which will eventually lead to all AVs when RVs are completely phased out. After the analysis of these specific policies, a general formulation regardless of the AV distribution will follow.

We study the effects of various parameters with different degrees of control feasibility on the capacity: (1) AV penetration rate p in traffic demand (external and essentially uncontrollable), (2) ε_1 and ε_2 (difficult to control but feasible with technological advancements); and (3) α_1 and α_2 (controllable). Namely, p is treated as an independent input variable; ε is a feature of the AV characteristics, also an independent input variable; and α is a controllable variable that can be used to optimize the capacity, condition on p and ε. For given p, ε_1 and ε_2, the capacity is achieved when the total flow reaches the maximum among the possible combinations of α_1 and α_2, denoted by Q_{case}^{max} with the subscript specifying the lane policy. The values of α_1 and α_2 corresponding to Q_{case}^{max} are denoted as α_1^* and α_2^*, respectively, referred to as the optimal solution to capacity. Additionally, among various p values, there is an optimum p that achieves the maximum level of capacity, which is referred to as the optimum capacity, Q_{case}^*. For each lane policy, we will first derive the relationship between the flow and key parameters, and then obtain Q_{case}^{max} and Q_{case}^*.

3.1.1 (A, R) policy

In the (A, R) case, $\alpha_1 = 1$ and $\alpha_2 = 0$ since AVs and RVs are segregated. Let $Q_{A,R}$ denote the total flow, $Q_{A,R}^{max}$ the capacity, and $Q_{A,R}^*$ the optimum capacity.
From the flow conservation of AVs and RVs, we have

$$\begin{cases} Q_{A,R} = q_1 + q_2 \\ pQ_{A,R} = q_1 \end{cases}. \tag{3-1}$$

Notice that q_1 and q_2 should not exceed their respective physical lane capacities, C_1 and C_2, which are given by Eqn. (2-4). By integrating $\alpha_1 = 1$ and $\alpha_2 = 0$ into Eqn. (2-4), we have the following constraints:

$$0 \le q_1 \le C_1 = \frac{C_0}{1-\varepsilon_1}, \tag{3-2a}$$

$$0 \le q_2 \le C_0 . \tag{3-2b}$$

By expressing q_1 and q_2 as functions of $Q_{A,R}$, the constraints in Eqn. (3-2) can be re-written as

$$pQ_{A,R} \le C_1 , \tag{3-3a}$$

$$(1-p)Q_{A,R} \le C_0 . \tag{3-3b}$$

Notice that since α_1 and α_2 are pre-determined in the (A, R) case, $Q_{A,R}^{max}$ equals to $Q_{A,R}$. Therefore, from Eqn. (3-3), we obtain $Q_{A,R}^{max}$ as follows:

$$\begin{cases} Q_{A,R}^{max} = \frac{C_0}{1-p} & , \; if \; 0 \le p < \frac{1}{2-\varepsilon_1} \\ Q_{A,R}^{max} = \frac{C_0}{p(1-\varepsilon_1)} & , \; if \; \frac{1}{2-\varepsilon_1} \le p \le 1 \end{cases}. \tag{3-4}$$

Or it can be written in a more general form:

$$Q_{A,R}^{max} = min\left(\frac{C_0}{1-p}, \frac{C_0}{p(1-\varepsilon_1)}\right). \tag{3-5}$$

Clearly, $Q_{A,R}^{max}$ varies with p, with the breakpoint, $p_{cric} = \frac{1}{2-\varepsilon_1}$. Notice that this breakpoint corresponds to the penetration rate, at which both lane 1 and lane 2 will reach their respective physical lane capacities. If $p < p_{cric}$, the first term of Eqn. (3-5) concerning lane 2 dominates. In this scenario, $Q_{A,R}^{max}$ is achieved when lane 2 reaches its physical lane capacity, C_0, while lane 1 is below its physical lane capacity ($q_1 < C_1$). $Q_{A,R}^{max}$ increases with p at an increasing rate, as shown in Fig. 3-2(a), since a higher p represents a higher utilization of lane 1 and thus, a higher $Q_{A,R}^{max}$. By contrast, if $p > p_{cric}$, the second term of Eqn. (3-5) concerning lane 1 dominates. In this scenario, $Q_{A,R}^{max}$ is achieved when lane 1 has reached its upper limit (C_1) but lane 2 is below its physical lane capacity, C_0. $Q_{A,R}^{max}$ decreases with p, as illustrated. This is because as p increases, lane 2 becomes more underutilized while the flow in lane 1 remains constant at C_1. Therefore, over the spectrum of $p \in [0,1]$, the maximum of $Q_{A,R}^{max}$ (i.e., $Q_{A,R}^{*}$), is achieved when $p = p_{cric}$, where both lanes have reached their respective physical lane capacities:

$$Q_{A,R}^{*} = max\left(Q_{A,R}^{max}\right) = C_0(Q_{A,R}^{max}) = C_0\left(1 + \frac{1}{1-\varepsilon_1}\right). \qquad \text{3-6)}$$

Notably, the result above suggests that if AVs and RVs are segregated, roadway capacity is underutilized unless $p = p_{cric}$. Therefore, to better utilize the capacity, the (M, R) policy should be considered if $p < p_{cric}$, to allow RVs to use lane 1, and the (A,M) policy if $p > p_{cric}$, to allow AVs to use lane 2. These will be discussed next.

It is also worth noting that, the AV efficiency gain in lane 1, ε_1, has an interesting effect on $Q_{A,R}^{max}$ according to Eqn. (3-5) and (3-6). Fig. 3-2(b) shows $Q_{A,R}^{max}$ plots and the corresponding $Q_{A,R}^{*}$ (captured by the dotted vertical lines) under various ε_1 values. Clearly, a greater ε_1 (a measure of the efficiency gain) leads to a greater p_{cric} and thus, $Q_{A,R}^{*}$. Since platooning is not allowed in lane 2, ε_2 does not have any impact on $Q_{A,R}^{max}$.

Fig. 3-2: (A, R) policy: (a) $Q_{A,R}^{max}$ under various p ($\varepsilon_1 = 0.5, \varepsilon_2 = 0.2$); (b) impacts of ε_1 on $Q_{A,R}^{max}$ and $Q_{A,R}^{*}$ ($\varepsilon_2 = 0.2$).

3.1.2 (M, R) policy

In the (M, R) case, a key feature is that α_1 is undetermined, though $\alpha_2 = 0$. Let $Q_{M,R}$ denote the total flow of the roadway, and similarly $Q_{M,R}^{max}$ denote the capacity for a given penetration rate p under the (M, R) policy. Based on the results of the (A, R) case above, we focus on the scenario of $p \leq p_{cric}$.

The flow conservation (in Eqn. (3-1) changes to the below:

$$\begin{cases} Q_{M,R} = q_1 + q_2 \\ pQ_{M,R} = \alpha_1 q_1 \end{cases}. \tag{3-7}$$

By reformulating Eqn. (3-7), we can derive the flows in the two lanes:

$$q_1 = \frac{pQ_{M,R}}{\alpha_1}, \tag{3-8a}$$

$$q_2 = Q_{M,R}\frac{\alpha_1 - p}{\alpha_1}. \tag{3-8b}$$

Similar to the (A, R) case, q_1 and q_2 should not exceed their physical lane capacities, C_1 and C_2, respectively, which are given by Eqn. (2-4). By plugging $\alpha_2 = 0$ into Eqn. (2-4), we obtain

$$q_1 \leq C_1 = \frac{C_0}{1-\alpha_1\varepsilon_1}, \tag{3-9a}$$

$$q_2 \leq C_0 . \tag{3-9b}$$

By expressing q_1 and q_2 as functions of $Q_{M,R}$ and integrating the premise of $p \leq p_{cric}$, the inequalities in (3-9) result in the following constraints for $Q_{M,R}$, respectively:

$$\begin{cases} Q_{M,R} \leq \frac{\alpha_1 C_0}{p(1-\alpha_1\varepsilon_1)} \ , \ if p \leq \alpha_1 \leq \frac{2p}{1+p\varepsilon_1} \\ Q_{M,R} \leq \frac{\alpha_1 C_0}{\alpha_1 - p} \quad \ , \ if \ \frac{2p}{1+p\varepsilon_1} \leq \alpha_1 \leq 1 \end{cases} . \tag{3-10}$$

Note that the lower bound for α_1 is p, obtained based on the assumption that $q_2 = 0$ (see Eqn. (3-8b)). Eqn. (3-10) can be written in a more general form:

$$Q_{M,R} \leq \min\left(\frac{\alpha_1 C_0}{p(1-\alpha_1\varepsilon_1)}, \frac{\alpha_1 C_0}{\alpha_1 - p}\right). \tag{3-11}$$

Eqn. (3-10) and (3-11) suggest that there exists a breakpoint of $\alpha_1 = \frac{2p}{1+p\varepsilon_1}$, at which both lane 1 and lane 2 reach their respective physical lane capacities (i.e., equalities in Eqn. (3-9)). If $\alpha_1 < \frac{2p}{1+p\varepsilon_1}$, lane 1 reaches the physical lane capacity but lane 2 is underused, and thus the first term (concerning lane 1) in Eqn. (3-11) determines $Q_{M,R}$. In this case, $Q_{M,R}$ increases with α_1 since C_1 increases with α_1 (Eqn. (3-9a)); see Fig. 3-3 (a). By contrast, if $\alpha_1 > \frac{2p}{1+p\varepsilon_1}$, the RV flow is constrained, and thus the second term (concerning lane 2) in Eqn. (3-11) determines $Q_{M,R}$. In this case, $Q_{M,R}$ decreases with α_1 since the greater the α_1, the smaller the available capacity for RVs ($= (1-\alpha_1)C_1 + C_0$), and thus the more constrained the RV flow.

Clearly, $Q_{M,R}$ achieves the maximum, $Q_{M,R}^{max}$, when $\alpha_1 = \frac{2p}{1+p\varepsilon_1}$; i.e., $\alpha_1^* = \frac{2p}{1+p\varepsilon_1}$ and $\alpha_2^* = 0$ by the nature of the policy. Then, $Q_{M,R}^{max}$ can be derived by plugging α_1^* into either expressions in Eqn. (3-11):

$$Q_{M,R}^{max} = \max(Q_{M,R}) = C_0 \frac{2}{1-p\varepsilon_1}. \tag{3-12}$$

Since we assume that $p \leq p_{cric}$, Eqn. (3-12) suggests that $Q_{M,R}^{max}$ increases with p; see Fig. 3-3(b). This is straightforward: if p is greater, more AVs can platoon in lane 1 (without exceeding its physical lane capacity C_1), resulting in more efficient use of the roadway. Notice that according to Eqn. (3-12), $Q_{M,R}^{max}$ reaches the maximum when $p = p_{cric}$:

$$Q_{M,R}^{*} = \max(Q_{M,R}^{max}) = C_0\left(1 + \frac{1}{1-\varepsilon_1}\right), \tag{3-13}$$

which converges to the optimum capacity of the (A, R) policy as expected; i.e., $Q_{M,R}^* = Q_{A,R}^*$. The effects of ε_1 and ε_2 are similar to the (A, R) policy. A comparison of the capacities under the (A, R) and (M, R) policies (i.e., $Q_{A,R}^{max}$ and $Q_{M,R}^{max}$) is shown in Fig. 3-3(c). The vertical difference between the two curves represents the additional capacity gain if one switches from the (A, R) to (M, R) policy. One can see that the two plots converge when $p = p_{cric}$. Note that if $p > p_{cric}$, the (M, R) policy is possible but it effectively becomes the (A, R) policy since lane 1 can be fully utilized by AVs before RVs in lane 2 reach C_0. In this case, lane 2 would be underutilized, and thus the (A, M) policy, presented below, is desirable to maximize the utilization of lane 2.

Fig. 3-3: (M, R) policy ($\varepsilon_1 = 0.5, \varepsilon_2 = 0.2$): (a) $Q_{M,R}$ under various α_1; (b) $Q_{A,R}^{max}$ under various p; (c) comparison of $Q_{A,R}^{max}$ and $Q_{M,R}^{max}$.

3.1.3 (A, M) policy

In this policy, lane 1 is dedicated to AVs only and lane 2 accommodates both AVs and RVs. Then the same principles of analysis as the (M, R) policy apply: we seek the optimal proportion for AVs in lane 2 (i.e., α_2^*) that maximizes the flow. In this case, $\alpha_1 = 1$ and α_2 is our decision variable. The total flow of the roadway is denoted by $Q_{A,M}$.

Let $Q_{A,M}$ denote the total flow. Based on the flow conservation, we obtain

$$\begin{cases} Q_{M,R} = q_1 + q_2 \\ pQ_{M,R} = \alpha_1 q_1 + \alpha_2 q_2 \end{cases} . \tag{3-14}$$

in which $\alpha_1 = 1$. By solving Eqn. (3-14) simultaneously, we derive q_1 and q_2:

$$q_1 = Q_{A,M} \frac{p - \alpha_2}{1 - \alpha_2}, \tag{3-15a}$$

$$q_2 = Q_{A,M} \frac{1 - p}{1 - \alpha_2}. \tag{3-15b}$$

Similar to the other policies, q_1 and q_2 should not exceed their respective physical lane capacities, C_1 and C_2, given by Eqn. (2-4). By plugging Eqn. (3-15) into Eqn. (2-4), we obtain

$$0 \le q_1 \le C_1 = \frac{C_0}{1 - 1 * \varepsilon_1}, \tag{3-16a}$$

$$0 \le q_2 \le C_2 = \frac{C_0}{1 - \alpha_2 \varepsilon_2}. \tag{3-16b}$$

By expressing q_1 and q_2 as functions of $Q_{A,M}$ and integrating them with the premise of $p > p_{cric}$, the inequalities in (3-16) result in the following constraints for $Q_{A,M}$ respectively:

$$\begin{cases} Q_{A,M} \le \frac{C_0(1-\alpha_2)}{(p-\alpha_2)(1-\varepsilon_1)}, \ if \ 0 \le \alpha_2 \le \frac{2p-(1+\varepsilon_1 p)}{1+\varepsilon_2 p-(\varepsilon_1+\varepsilon_2)} \\ Q_{A,M} \le \frac{C_0(1-\alpha_2)}{(1-p)(1-\alpha_2\varepsilon_2)}, \ if \ \frac{2p-(1+\varepsilon_1 p)}{1+\varepsilon_2 p-(\varepsilon_1+\varepsilon_2)} \le \alpha_2 \le \mathrm{p} \end{cases} . \tag{3-17}$$

More generally,

$$Q_{A,M} \le \min\left(\frac{C_0(1-\alpha_2)}{(p-\alpha_2)(1-\varepsilon_1)}, \frac{C_0(1-\alpha_2)}{(1-p)(1-\alpha_2\varepsilon_2)}\right). \tag{3-18}$$

The upper bound for α_2, p, is obtained by setting $q_1 = 0$ in Eqn. (3-15a).

The interpretation of these constraints are similar to the (M, R) policy. Specifically, the breakpoint of $\alpha_2 = \frac{2p-(1+\varepsilon_1 p)}{1+\varepsilon_2 p-(\varepsilon_1+\varepsilon_2)}$ corresponds to the AV proportion in lane 2 that both lanes achieve their respective physical lane capacities (i.e., equalities in Eqn. (3-16) and (3-17)). If α_2 is smaller than this breakpoint, lane 1 is underused. Thus, the first term concerning lane 1 in Eqn. (3-18) determines $Q_{A,M}$, in which case $Q_{A,M}$ increases with α_2 since C_2 increases with α_2; see Fig. 3-4 (a). Otherwise, AV flow in lane 2 is constrained, and the second term determines $Q_{A,M}$. In this case, $Q_{A,M}$ decreases with α_2 since the greater the α_2, the more constrained the AV flow.

Fig. 3-4: (A,M) policy ($\varepsilon_1 = 0.5, \varepsilon_2 = 0.2$): (a) $Q_{A,M}$ under various α_2; (b) $Q_{A,M}^{max}$ under various p; (c) $Q_{A,M}^{max}$ under various ε_2;(c) comparison of $Q_{A,R}^{max}$, $Q_{M,R}^{max}$, and $Q_{A,M}^{max}$.

As in the (M, R) policy, the capacity under this policy, $Q_{A,M}^{max}$, occurs at the breakpoint; i.e., $\alpha_2^* = \frac{2p-(1+\varepsilon_1 p)}{1+\varepsilon_2 p-(\varepsilon_1+\varepsilon_2)}$ (and $\alpha_1^* = 1$). Then, $Q_{A,M}^{max}$ is derived as:

$$Q_{A,M}^{max} = \max(Q_{A,M}) = \frac{C_0(2-\varepsilon_1-\varepsilon_2)}{(1-\varepsilon_1)(1-p\varepsilon_2)}. \quad (3\text{-}19)$$

Since we expect that $0 \le \varepsilon_1 < 1$ and $0 \le \varepsilon_2 < 1$, Eqn. (3-19) suggests that $Q_{A,M}^{max}$ increases with p; see Fig. 3-4 (b). This is expected: as p increases, more AVs will platoon in lane 2, resulting in more efficient use of the roadway. Thus, $Q_{M,R}^{*}$ is achieved when $p = 1$ (i.e., all AVs):

$$Q_{M,R}^{*} = \max(Q_{A,M}^{max}) = \frac{C_0(2-\varepsilon_1-\varepsilon_2)}{(1-\varepsilon_1)(1-\varepsilon_1)} = \frac{C_0}{1-\varepsilon_1} + \frac{C_0}{1-\varepsilon_2}. \quad (3\text{-}20)$$

From Eqn. (3-20), it is also intuitive that $Q_{A,M}^{max}$ increases with ε_1 and ε_2, the measurement of AV efficiency in lanes 1 and 2, respectively. Specifically, the impacts of ε_1 are the same with the (M, R) policy, but it is worth noting that unlike the (M, R) policy, ε_2 now plays an important role in $Q_{A,M}^{max}$ and $Q_{M,R}^{*}$ because platooning in lane 2 results in more efficient roadway usage; see Fig. 3-4(c). A comparison of $Q_{A,M}^{max}$ to the capacity under the (A, R) policy ($Q_{A,R}^{max}$) is shown in Fig. 3-4(d). It is clear that allowing AV platooning in lane 2 substantially increases the capacity (from the red to the blue curve). Note that if $p \le p_{cric}$, the (A, M) policy effectively becomes the (A, R) policy since, with sufficient demand, RVs in lane 2 will reach C_0 before AVs fully utilize lane 1. As discussed, the (M, R) policy is desired in this case to further increase the capacity. Table 1 summarizes the capacity states for the three policies.

Table 1: Summary of capacity formulation for three lane policies

Policy	$0 < p \le p_{cric}$			$p_{cric} < p \le 1$		
(A,R)	$\alpha_1^* = 1$	$\alpha_2^* = 0$	$Q_{A,R}^{max} = \frac{C_0}{1-p}$	$\alpha_1^* = 1$	$\alpha_2^* = 0$	$Q_{A,R}^{max} = \frac{C_0}{p(1-\varepsilon_1)}$
(M,R)	$\alpha_1^* = \frac{2p}{1+p\varepsilon_1}$	$\alpha_2^* = 0$	$Q_{M,R}^{max} = C_0 \frac{2}{1-p\varepsilon_1}$	Same as (A, R) case		
(A,M)	Same as (A, R) case			$\alpha_1^* = 1$	$\alpha_2^* = \frac{2p-(1+\varepsilon_1 p)}{1+\varepsilon_2 p-(\varepsilon_1+\varepsilon_2)}$	$Q_{A,M}^{max} = \frac{C_0(2-\varepsilon_1-\varepsilon_2)}{(1-\varepsilon_1)(1-p\varepsilon_2)}$

The capacity formulations for the three policies provide a useful insight for implementation. Particularly, they show the possible and the most efficient solutions for a given demand level (Q) and penetration rate (p). This is illustrated in Fig. 3-5. Specifically, under small penetration rates ($p \le p_{cric}$), all three policies are possible if $Q \le Q_{A,R}^{max}$ (see the pink region); while only the (M, R) policy should be considered if $Q > Q_{A,R}^{max}$ (the purple region). Otherwise, the flow will be constrained at $Q_{A,R}^{max}$. With greater penetration rates ($p > p_{cric}$), all three policies are possible if $Q \le Q_{A,R}^{max}$ (the blue region), but only the (A, M) policy is desirable if $Q > Q_{A,R}^{max}$ (the green region). These regions will be fully investigated in the next section.

Fig. 3-5: Feasible policies under various demand.

Note that these conclusions are based on our premise that lane 1 is more efficient than lane 2 ($\varepsilon_1 > \varepsilon_2$). If we assume the opposite ($\varepsilon_1 < \varepsilon_2$), the results will be reversed and symmetric. The general formulation relaxing this assumption will be provided in the next section.

3.2 General formulation for two-lane highway

The capacity formulations in Section 3.1 provide insight into which policy should be adopted to achieve the highest capacity if the overall demand is high. However, if the demand is sufficiently low, multiple policies may be feasible. This section presents a general framework, inclusive of all the lane policies considered in this paper, to determine valid domains of different lane policies and more generally, lane distributions of AVs with respect to demand. Specifically, in our general framework, AVs and RVs can use both lanes, and thus, the AV proportions, α_1 and α_2, are our decision variables. Additionally, we eliminate the assumption that lane 1 is more efficient than lane 2; namely, now it is possible to have $\varepsilon_1 \geq \varepsilon_2$ or $\varepsilon_1 \leq \varepsilon_2$.

In the general formulation, flow conservation can be written as follows:

$$\begin{cases} Q = q_1 + q_2 \\ pQ = \alpha_1 q_1 + \alpha_2 q_2 \end{cases}, \tag{3-21}$$

where Q denotes the total flow, and q_i denotes the flow in lane i, which should not exceed the physical lane capacity; i.e.,

$$0 \leq q_1 \leq C_1 , \tag{3-22a}$$

$$0 \leq q_2 \leq C_2 , \tag{3-22b}$$

where C_i depends on α_i and ε_i according to Eqn. (2-4). Additionally, the AV proportions have natural physical bounds:

$$0 \leq \alpha_1 \leq 1 , \tag{3-23a}$$

$$0 \leq \alpha_2 \leq 1 . \tag{3-23b}$$

By solving the equations in (3-21) simultaneously, we express q_1 and q_2 as:

$$\begin{cases} q_1 = \frac{(p-\alpha_2)Q}{\alpha_1-\alpha_2} \\ q_2 = \frac{(\alpha_1-p)Q}{\alpha_1-\alpha_2} \end{cases}, \; if \; \alpha_1 \neq \alpha_2 , \tag{3-24a}$$

and

$$q_1 + q_2 = Q, \; if \; \alpha_1 = \alpha_2 = p . \tag{3-24b}$$

Note that Eqn. (3-24b) denotes a special value set of (α_1, α_2), for which q_1 and q_2 have infinite solutions as long as their sum is Q and they satisfy (3-21).

The formulations in Eqn. (3-24) are subject to the constraints in Eqn. (3-22) and (3-23), which will define the valid domains for (α_1, α_2). Particularly, by integrating Eqn. (3-24), (3-22), and the physical lane capacity function (2-4), we obtain the formulations of α_1 and α_2 constraints in a simple but noble structure. Specifically, if $\alpha_1 \leq p$,

$$(\alpha_1 - \xi_1)(\alpha_2 - \xi_2) \geq \xi_3, \; if \; 0 \leq \alpha_1 \leq p , \tag{3-25a}$$

$$(\alpha_1 - \xi_4)(\alpha_2 - \xi_5) \leq \xi_6, \; if \; 0 \leq \alpha_1 \leq p . \tag{3-25b}$$

Or if $\alpha_1 > p$

$$(\alpha_1 - \xi_1)(\alpha_2 - \xi_2) \leq \xi_3, \; if \; p < \alpha_1 \leq 1 , \tag{3-26a}$$

$$(\alpha_1 - \xi_4)(\alpha_2 - \xi_5) \geq \xi_6, \; if \; p < \alpha_1 \leq 1 . \tag{3-26b}$$

where

$$\xi_1 = \frac{1}{\varepsilon_1} - \frac{C_0}{Q\varepsilon_1},$$
$$\xi_2 = p + \frac{C_0}{Q\varepsilon_1},$$
$$\xi_3 = \frac{C_0(Q(1-p\varepsilon_1)-C_0)}{Q^2{\varepsilon_1}^2},$$
$$\xi_4 = p + \frac{C_0}{Q\varepsilon_2},$$
$$\xi_5 = \frac{1}{\varepsilon_2} - \frac{C_0}{Q\varepsilon_2},$$
$$\xi_6 = \frac{C_0(Q(1-p\varepsilon_2)-C_0)}{Q^2{\varepsilon_2}^2}.$$

The valid domains of (α_1, α_2) are the regions that satisfy Eqn. (3-23), (3-25), and (3-26); see the shaded regions in Fig. 3-6, for example, with $p = 0.4$ ($p < p_{cric}$), $\varepsilon_1 = 0.5, \varepsilon_2 = 0.2$. In these plots, the two boundaries (red and green dashed lines) denote the instances where the equalities to the upper bounds are achieved in Eqn. (3-22): the red dashed boundary from $q_1 = C_1$ and the green dashed boundary from $q_2 = C_2$. (Note that if the demand is extremely low (e.g., $Q \ll C_0$), it becomes impossible to achieve the equalities, and the boundaries will originate from the natural physical boundaries in Eqn. (3-23)). The figure also shows contours of q_1 (in relative scale of C_0), corresponding to a set of linear relationship between α_1 and α_2, derived from the first formulation in (3-24a):

$$\alpha_2 = \frac{pQ}{Q-q_1} - \frac{q_1}{Q-q_1}\alpha_1 .$$

Specifically, the valid domains are colored according to the q_1/C_0 values, and the black lines denote contours for some specific q_1 values; see Fig. 3-6(a). One can see that q_1 increases in the clock-wise direction as the color transitions from dark blue to light orange. A similar feature can be obtained for q_2, where q_2 increases in the counter-clockwise direction, opposite to q_1.

Five features of valid domains are worth noting from the figure.

R1: The valid domains fall in two regions, upper left and lower right since α_1 and α_2 should satisfy the following conditions according to Eqn. (3-24) to assure that $q_1, q_2 \geq 0$:

$$\alpha_1 \geq p \text{ and } \alpha_2 \leq p\,, \tag{3-27a}$$

Or

$$\alpha_1 \leq p \text{ and } \alpha_2 \geq p\,. \tag{3-27b}$$

R2: The valid domains shrink as Q increases; for example, see Fig. 3-6(a-c) for valid domains for three demand levels in the relative scale of $2C_0$, representing low, medium, and high demand. This implies that the feasible solution set (i.e., combinations of α_1 and α_2 and consequently q_1 and q_2) decreases as the demand increases. Notably, if the demand is sufficiently high, the solution reduces to a single point, suggesting that there is a unique solution.

R3: The valid domains vary with the penetration rate p. Fig. 3-7 shows the valid domains with the same set of demand levels (and ε_1 and ε_2) but with $p = 0.7 > p_{cric}$. One can see that the valid domains shift towards the upper right, closer to the physical boundary where $\alpha_1 = 1$ and $\alpha_2 = 1$. This is expected because with a large penetration rate, the chance of having a lane fully filled by AVs is greater.

R4: For given p, ε_1 and ε_2, feasible policies vary with the demand level, which are denoted by the bold lines along the boarders. For example, in low demand (Fig. 3-6(a)) six policies are possible. In medium demand (Fig. 3-6(b)), the feasible policies reduce to (M, R) and (R, M), and in high demand (Fig. 3-6(c)), only (M, R) is possible. Obviously, changes of feasible policies also depend on p. One can see the differences in feasible policies between Fig. 3-6 and Fig. 3-7.

R5: The solution to the maximum possible demand (i.e., the level of capacity) varies with p and the relationship between ε_1 and ε_2. Each dot in Fig. 3-8(a) denotes the solution for a given p value in the case of $\varepsilon_1 > \varepsilon_2$, which is essentially obtained from the formulation in Table 1. Clearly, when $\leq p_{cric}$, the solutions are found on the bottom border ($\alpha_2 = 0$), corresponding to the (M, R) policy, and move towards right as p increases. When $p > p_{cric}$, the solutions are found on the right boarder ($\alpha_1 = 1$), corresponding to the (A, M) policy, and move towards the top boarder. Interestingly, when the relationship between ε_1 and ε_2 is reversed (i.e., $\varepsilon_1 < \varepsilon_2$), the solutions occur on the left and then move to the top boarder as p increases (see Fig. 3-8b), which is because the roles of ε_1 and ε_2 are symmetric. If $\varepsilon_1 = \varepsilon_2$ (i.e., both lanes have the same AV gain), a unique solution is found only when $p = 0 \; or \; 1$. Otherwise, there are infinite solutions of α_1 and α_2 that satisfy the relationship in Eqn. (3-21) and lead to the same capacity. This is illustrated in Fig. 3-8(c). Rather than a unique solution, a contour line is obtained for a given p value ($0 < p < 1$). This contour line is obtained by plugging the physical lane capacity function Eqn. (2-4) into Eqn. (3-24) and solving the equation, which is given as follows:

$$\left(\alpha_1 - \frac{1+p\varepsilon}{2\varepsilon}\right)\left(\alpha_2 - \frac{1+p\varepsilon}{2\varepsilon}\right) = \left(\frac{1-p}{2}\right)^2, \tag{3-28}$$

where $\varepsilon = \varepsilon_1 = \varepsilon_2$.

Fig. 3-6: Valid domain for $p < p_{cric}$ ($p = 0.4, \varepsilon_1 = 0.5, \varepsilon_2 = 0.2$): (a) low demand ($\frac{Q}{2C_0} = 0.75$); (a) medium demand ($\frac{Q}{2C_0} = 0.94$); (c) high demand ($\frac{Q}{2C_0} = 1.125$).

Fig. 3-7: Valid domain for $p > p_{cric}$ ($p = 0.7, \varepsilon_1 = 0.5, \varepsilon_2 = 0.2$): (a) low demand ($\frac{Q}{2C_0} = 0.75$); (a) medium demand ($\frac{Q}{2C_0} = 0.94$); (c) high demand ($\frac{Q}{2C_0} = 1.125$).

Fig. 3-8 Solution to operational capacity under various p: (a) scenario of $\varepsilon_1 > \varepsilon_2$; (a) scenario of $\varepsilon_1 < \varepsilon_2$; (c) scenario of $\varepsilon_1 = \varepsilon_2$.

4 Multi-lane Highway

This section aims to expand the general framework for a two-lane highway to the general K-lane case, with $K > 1$. We assume that for lane i, the efficiency gain with AV platooning is denoted by ε_i, the flow q_i, and the AV proportion α_i. The total flow is denoted by Q_K. To facilitate a simpler formulation, we assume that the lanes are numbered according to their ε_i values in the descending order; i.e., $\varepsilon_1 \geq \varepsilon_2 \geq \varepsilon_3 \geq \cdots \geq \varepsilon_K$, but this is not required to obtain the following results.

Similar to the two-lane case, from the flow conservation of AVs and RVs Eqn. (3-21), we have:

$$\begin{cases} Q_K = \sum_{i=1}^{K} q_i \\ pQ_K = \sum_{i=1}^{K} \alpha_i q_i \end{cases}, \tag{4-1}$$

where the flow and AV proportions are constrained by the physical lane capacities:

$$0 \le q_i \le C_i\,,\ \ i = 1,2,\cdots,K \tag{4-2a}$$

$$0 \le \alpha_i \le 1\,,\ \ i = 1,2,\cdots,K \tag{4-2b}$$

And C_i is given by the physical lane capacity function in Eqn. (2-4). We have proved that to maximize the overall capacity, each lane should reach its physical lane capacity; i.e., $q_i = C_i$; see the proof in Appendix A. Thus, from Eqn. (4-1), we can reformulate the flow conservation and the capacity of K-lane, Q_K^{max}, as follows:

$$Q_K^{max} = \max_{Q_K}\left\{ Q_K \,\middle|\, \sum_{i=1}^{K} \frac{p-\alpha_i}{1-\alpha_i\varepsilon_i} = 0, 0 \le p \le 1, 0 \le \varepsilon_i < 1, 0 \le \alpha_i \le 1, \forall\, 0 \le i \le K, K > 1 \right\}. \tag{4-3}$$

It can be proved (the proof is in the Appendix B) that the solution of $(\alpha_1, \alpha_2,\ldots, \alpha_K)$ to Q_K^{max} is below:

$$\alpha_i = 1, \forall\, 1 \le i \le J-1, if\ 1 < J < K\,, \tag{4-4a}$$

$$\alpha_i = 0, \forall\, J+1 \le i \le K, if\ 1 < J < K\,, \tag{4-4b}$$

$$0 \le \alpha_J \le 1, \tag{4-4c}$$

if p falls in the following range:

$$\frac{\phi_{J-1}}{\phi_{J-1}+K-J+1} \le p \le \frac{\phi_J}{K-J+\phi_J}, \tag{4-5}$$

where ϕ_{J-1} measures the physical lane capacity sum of lanes 1 to $J-1$ (in relative scale of C_0), given as below:

$$\phi_{J-1} = \sum_{i=1}^{J-1} \frac{1}{1-\varepsilon_i}\,. \tag{4-6}$$

The solution in Eqn. (4-4) and (4-5) implies that lanes 1 to $J-1$ only have AVs, lanes $J+1$ to K only have RVs, and lane J has mixed traffic. The physical mechanism behind this solution is straightforward: to achieve the maximum flow (i.e., capacity), AVs should fill the most efficient lane first and gradually move to the less efficient lanes. This is consistent with our results for the two-lane highway scenario.

From Eqn. (4-3) and (4-4), we can derive α_J and Q_K^{max}:

$$\alpha_J = \frac{p(K-J+1+\phi_{J-1})-\phi_{J-1}}{1-\varepsilon_J\phi_{J-1}+p\varepsilon_J(K-J+\phi_{J-1})}\,. \tag{4-7}$$

$$Q_K^{max} = C_0\left(\sum_{i=1}^{J-1}\frac{1}{1-\varepsilon_i} + \frac{1}{1-\alpha_J\varepsilon_J} + K - J\right) = C_0\frac{K-J+1+(1-\varepsilon_J)\phi_{J-1}}{1-p\varepsilon_J}\,. \tag{4-8}$$

The derivative of α_J with respect to p is always positive, suggesting that α_J increases with p. This is also straightforward: as p increases, the AV proportion in lane J increases until it reaches the maximum ($\alpha_J = 1$). Thereafter, AVs will start to use lane $J+1$. More importantly, Q_K^{max} increases with p since $\phi_{J-1} > J-1$ always holds (see Eqn. (4-6)) and thus the numerator in Eqn. (4-8) is always positive. From the bounds of p in Eqn. (4-5), the bounds of Q_K^{max} can be derived:

$$C_0\big(K-J+1+\phi_{J-1}\big) \le Q_K^{max} \le C_0\big(K-J+\phi_J\big)\,. \tag{4-9}$$

The lower and upper bounds of Q_K^{max} are achieved when $\alpha_J = 0$ and $\alpha_J = 1$, respectively. In the full spectrum of $p \in [0,1]$, the maximum of Q_K^{max}, namely the optimum capacity, Q_K^*, is achieved when $p = 1$, and all lanes will be filled by AVs:

$$Q_K^* = C_0 \phi_K \,. \tag{4-10}$$

For generic traffic demand, we can derive valid domains of α_i's in a way similar to the two-lane highway scenario by addressing the constraints in Eqn. (4-2). However, one can expect K decision variables, and exact solutions to the bounds will be complex. Nevertheless, numerical solutions can be obtained.

5 Potential correlation between ε and α

In formulating the capacity so far, ε and α are treated as independent based on the assumption that the platoon size, n, is fixed. In this section, this assumption is relaxed since n could increase with α (e.g., drivers tend to form longer platoons when they see more AVs around). More generally, we investigate potential correlation between ε and α and its effect on capacity and optimal lane policies. In this case, ε is dynamic, increasing with α, and the physical lane capacity function (Eqn. (2-4)) also becomes dynamic. However, traffic would eventually reach an equilibrium. Theoretically ε can be negatively correlated with α too. Either way, the formulations for flow conservation and constraints still hold, but the optimal lane policy could be different. A more in-depth investigation follows.

For the correlation case, we consider a two-lane highway. Assume that ε_1 (ε_2) is a function of α_1 (α_2); i.e., $\varepsilon_1 = E_1(\alpha_1)$ and $\varepsilon_2 = E_2(\alpha_2)$. It can be proved that when a two-lane highway reaches its optimal capacity, all its lanes should have reached their physical lane capacity (see proof in Appendix C part I). Let $\tilde{Q}$ denote the total flow when both lanes reach their respective physical lane capacities. Then, we can formulate an optimization problem to maximize $\tilde{Q}$ (objective function) with respect to α_1 and α_2 (decision variables):

$$\tilde{Q}(\alpha_1, \alpha_2) = C_0 \frac{1}{1-\alpha_1 E_1} + C_0 \frac{1}{1-\alpha_2 E_2}\,, \tag{5-1}$$

subject to the flow conservation and physical boundaries of α_1, α_2, and p:

$$C_0 \frac{p-\alpha_1}{1-\alpha_1 E_1} + C_0 \frac{p-\alpha_2}{1-\alpha_2 E_2} = 0, \tag{5-2}$$

$$0 \le \alpha_1 \le 1, \tag{5-3a}$$

$$0 \le \alpha_2 \le 1, \tag{5-3b}$$

$$0 \le p \le 1. \tag{5-3c}$$

Note that this is not a convex optimization problem because the constraint for the flow conservation (Eqn. (5-2)) is not necessarily convex (notice that E_1 and E_2 are functions of α_1 and α_2). Thus, this problem can be solved using a heuristic algorithm (e.g., genetic algorithms). To gain better insight into the effect of α_1 and α_2 on $\tilde{Q}$, we perform a marginal analysis. The results are complex due to the complex feasible domains of α_1 and α_2 and thus are presented in Appendix C part II. Instead, we present the result of a numerical experiment by assuming some typical functions for $E_i(\alpha_i)$ to obtain some insight.

We consider three different typical functions for $E_i(\alpha_i)$ and obtain the feasible solution domain of $\tilde{Q}$ for each. Specifically, we consider (i) linear function, $E_i(\alpha_i) = \beta_0 + \beta_1 \alpha_i$; (ii) convex function, $E_i(\alpha_i) = \beta_0 + \beta_1 {\alpha_i}^2$; and (iii) concave function, $E_i(\alpha_i) = \beta_0 + \beta_1 * \sqrt{\alpha_i}$; see Fig. 5-1. We also assume that $E_1(\alpha_1)$ and $E_2(\alpha_2)$ have the same functional form with some vertical shift: $E_2(\alpha_2)$ is shifted below $E_1(\alpha_1)$. Thus, it is assumed that lane 1 has larger efficiency gains for any given AV proportion. The results are illustrated in Fig. 5-2. For each function, two levels of p are considered, representing small and large penetration rates: $p = 0.2$ and $p = 0.8$, respectively. Note that in each plot, the solid black line represents the feasible domain, and the black dot denotes the optimal solution. For all three functions considered, α_2 decreases with α_1 in the feasible domains. For the small p value (Fig. 5-2(a, c, e)), the optimum of $\tilde{Q}$ is achieved in the bottom border, where $\alpha_2 = 0$ and $\alpha_1 > 0$, suggesting that all AVs are allocated in the more efficient lane (lane 1) and that lane 1 is partially filled with AVs due to the small penetration. For the large p value (Fig. 5-2(b, d, f)), the optimal solution is found on the right boarder, where $\alpha_2 > 0$ and $\alpha_1 = 1$, suggesting that lane 1 is fully filled by AVs while lane 2 has mixed traffic. Clearly, the optimal AV distribution strategies here are consistent with the case of independent α_i and ε_i, though the capacity values are different. Obviously, the multi-lane case will have more complex feasible domains and optimal solutions. However, the finding for the two-lane case can be generalized, and we conjecture that the optimal strategies will be consistent with the independent case. To be succinct, formulation for the multi-lane case is omitted.

Fig. 5-1: Four functions for $E_i(\alpha_i)$: (a) linear function ($E_1(\alpha_1) = 0.3 + 0.4\alpha_1$, $E_2(\alpha_2) = 0.4\alpha_2$); (b) convex function ($E_1(\alpha_1) = 0.3 + 0.4{\alpha_1}^2$, $E_2(\alpha_2) = 0.4{\alpha_2}^2$); (c) concave function ($E_1(\alpha_1) = 0.3 + 0.4\sqrt{\alpha_1}$, $E_2(\alpha_2) = 0.4\sqrt{\alpha_2}$).

Fig. 5-2: Feasible domains and optimal solution with correlated α_i and $E_i(\alpha_i)$: (a-b) linear function ($E_1(\alpha_1) = 0.3 + 0.4\alpha_1$, $E_2(\alpha_2) = 0.4\alpha_2$); (c-d) convex function ($E_1(\alpha_1) = 0.3 + 0.4{\alpha_1}^2$, $E_2(\alpha_2) = 0.4{\alpha_2}^2$); (e-f) concave function ($E_1(\alpha_1) = 0.3 + 0.4\sqrt{\alpha_1}$, $E_2(\alpha_2) = 0.4\sqrt{\alpha_2}$); $p = 0.2$ for (a, c, e) and $p = 0.8$ for (b, d, f).

6 Conclusion and Discussion

In this study, we developed a general theoretical framework to study how the macroscopic capacity in equilibrium traffic will change with the introduction of AVs. We first derived the formulation for the physical lane capacity (independent of the AV penetration rate, p) of a single lane facility considering micro/mesoscopic characteristics of RVs and AVs, including the platoon size and spacing characteristics. Particularly, efficiency gain via AV platooning is expressed as a parameter in terms of spacing characteristics of different vehicle types and platoon positions, which establishes a clear connection between microscopic vehicle characteristics and macroscopic roadway capacity. Based on the formulation for a single-lane facility, we formulated the capacities of a two-lane highway, incorporating p, for different lane policies to accommodate AVs, from segregated AV and/or RV lanes to mixed-use lanes. We found that strict segregation of AVs and RVs ((A, R) policy) can lead to lower capacity and that mixed-use ((M, R) and (A, M)) policies can realize higher capacities. Based on the analytical formulation, we determined the optimal policy and AV distribution, depending on the AV penetration rate.

We further developed a general formulation, inclusive of all the lane policies considered in this paper, to determine how AVs should be distributed across lanes, given traffic demand, AV penetration rate, and AV efficiency gain. The analytical formulation offered important insight into valid domains for AV distributions. We found that as demand increases, feasible domains of AV distributions shrink and eventually converge to a unique solution or a contour line that depends on the AV penetration and efficiency gain parameters. Lastly, we extended our formulation to the general multi-lane highway. It was found that to make the best utilization of roadway efficiency and thus achieve the highest capacity, AVs should use the most efficient lane(s) (i.e., with the largest AV gain) to the maximum possible extent and then move to the less efficient lanes.

We also explored the impact of correlated ε and α on capacity and optimal solution. Our numerical experiments via three typical functions suggest that for a two-lane highway the main conclusions are consistent with the case of independent ε and α. We conjecture that this finding can be generalized for multi-lane cases, but a more comprehensive investigation is needed in the future. Particularly, future research is needed to (1) understand the potential correlations based on empirical tests, and (2) integrate the correlations in the capacity formulation and determine the new optimal lane management policy.

Several extensions to the present study are desired in the future. Our formulations were developed based on average speed and spacing characteristics. A more systematic investigation of heterogeneous driver/vehicle characteristics is necessary to fully understand the impact of AVs on traffic capacity, such as different vehicle length, spacing preference, and platoon size. ε can be modified to capture these effects, such that it varies across drivers (or driver types). Then, the overall effect can be assessed by considering the distributions of driver/vehicle characteristics, as illustrated by Treiber and Kesting (Treiber and Kesting, 2013) for regular vehicles, for example. Moreover, although our capacity formulations capture some key microscopic characteristics of vehicle spacing, it does not capture the details of a platooning process. When an AV enters or exits a lane, it will likely require an extra spacing (Milanes et al., 2014; Nowakowski et al., 2016) or create a void ahead as in regular traffic (Laval and Daganzo, 2006), which can compromise the capacity. The net effect of each maneuver can potentially be reflected in ε. Another issue is traffic instability during platoon splitting and merging. Although successful platooning is presumed to have string stability (Milanes et al., 2014), it is unclear how traffic instability will involve when the AVs interact with RVs.

Additionally, some assumptions made in this study could be violated, and we caution against generalizing the results. Firstly, we treat AV penetration rate, p, as an external variable that remains consistent throughout the system. This assumption holds if FIFO can be maintained, which is reasonable if traffic is freely flowing (with similar desirable speed) and no prioritization scheme is applied. However, the FIFO principle may be violated upstream of a control region, particularly in (A, R) policy, where different traffic regimes can arise in different lanes (e.g., when congestion starts to build up), and vehicles self-organize. In that case, the traffic composition in the controlled region may differ significantly from what is expected from the AV penetration rate. Secondly, our formulation uses a deterministic and static p value. In reality, p may vary over time and location, which can affect optimal lane policies. Capacity formulation considering stochastic and/or dynamic p is left for future research.

Each problem discussed above presents an important and very challenging research topic. Extensive future research is needed, building on empirical studies of AV behavior and the interactions between AVs and RVs to better understand the mixed traffic. Nevertheless, this paper provides an explicit framework to link these microscopic characteristics to the macroscopic operational capacity.

Acknowledgement

This research was sponsored by the National Science Foundation through Award CMMI 1536599 and the Wisconsin Traffic Operations and Safety (TOPS) Laboratory at the University of Wisconsin-Madison.

References

Ahn, S., Cassidy, M.J., Laval, J., 2004. Verification of a simplified car-following theory. Transp. Res. Part B Methodol. 38, 431–440. doi:10.1016/S0191-2615(03)00074-2

Bu, F., Tan, H.-S., Huang, J., 2010. Design and field testing of a cooperative adaptive cruise control system. Am. Control Conf. (ACC), 2010. doi:10.1109/ACC.2010.5531155

Jerath, K., Brennan, S.N., 2012. Analytical prediction of self-organized traffic jams as a function of increasing acc penetration. Intell. Transp. Syst. IEEE Trans. doi:10.1109/TITS.2012.2217742

Kesting, A., Treiber, M., Helbing, D., 2010. Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity. Philos. Trans. R. Soc. London A Math. Phys. Eng. Sci. 368, 4585–4605. doi:10.1098/rsta.2010.0084

Kesting, A., Treiber, M., Schönhof, M., Helbing, D., 2008. Adaptive cruise control design for active congestion avoidance. Transp. Res. Part C Emerg. Technol. 16, 668–683. doi:10.1016/j.trc.2007.12.004

Kesting, A., Treiber, M., Schönhof, M., Helbing, D., 2007. Extending adaptive cruise control to adaptive driving strategies. Transp. Res. Rec. 2000, 16–24. doi:10.3141/2000-03

Laval, J.A., Daganzo, C.F., 2006. Lane-changing in traffic streams. Transp. Res. Part B Methodol. 40, 251–264. doi:10.1016/j.trb.2005.04.003

Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves. i. flood movement in long rivers. Proc. R. Soc. A Math. Phys. Eng. Sci. 229, 281–316. doi:10.1098/rspa.1955.0088

Lin, P.-S., Wang, Z., 2013. Impact of automated vehicles on highway safety and operations.

Litman, T., 2015. Autonomous vehicle implementation predictions implications for transport planning, in: Transportation Research Board 94th Annual Meeting. Washington DC.

Milanés, V., Shladover, S.E., 2014. Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. Transp. Res. Part C Emerg. Technol. 48, 285–300. doi:10.1016/j.trc.2014.09.001

Milanés, V., Shladover, S.E., Spring, J., Nowakowski, C., 2014. Cooperative adaptive cruise control in real traffic situations. IEEE Trans. Intell. Transp. Syst. 15, 296–305.

Milanes, V., Shladover, S.E., Spring, J., Nowakowski, C., Kawazoe, H., Nakamura, M., 2014. Cooperative adaptive cruise control in real traffic situations. IEEE Trans. Intell. Transp. Syst. 15, 296–305. doi:10.1109/TITS.2013.2278494

Newell, G.F., 2002. A simplified car-following theory: a lower order model. Transp. Res. Part B Methodol. 36, 195–205. doi:10.1016/S0191-2615(00)00044-8

NHTSA, 2013. U.s. department of transportation releases policy on automated vehicle development [WWW Document]. URL http://www.nhtsa.gov/About+NHTSA/Press+Releases/U.S.+Department+of+Transportation+Releases+Policy+on+Automated+Vehicle+Development

Nowakowski, C., Author, C., Station, R.F., Thompson, D., Lu, X., 2016. Operational concepts for truck cooperative adaptive cruise control (cacc) maneuvers, in: Transportation Research Board 95th Annual Meeting. pp. 1–16.

Ploeg, J., Scheepers, B.T.M., van Nunen, E., van de Wouw, N., Nijmeijer, H., 2011. Design and experimental evaluation of cooperative adaptive cruise control. 4th Int. IEEE Conf. Intell. Transp. Syst. (ITSC), 2011. doi:10.1109/ITSC.2011.6082981

Richards, P.I., 1956. Shock waves on the highway. Oper. Res. 4, 42–51. doi:10.2307/167515
Shladover, S.E., Nowakowski, C., O'Connell, J., Cody, D., 2010. Cooperative adaptive cruise control: driver driver selection selection of carcar-following gaps, in: 17th ITS World Congress. Busan , South Korea.
Shladover, S.E., Su, D., Lu, X.-Y., 2012. Impacts of cooperative adaptive cruise control on freeway traffic flow. Transp. Res. Rec. J. Transp. Res. Board 2324, 63–70. doi:10.3141/2324-08
Swaroop, D., Hendrick, J.K., Chien, C.C., Ioannou, P., 1994. A comparision of spacing and headway control laws for automatically controlled vehicles. Veh. Syst. Dyn. 23, 579–625. doi:10.1080/00423119408969077
Talebpour, A., Mahmassani, H.S., 2016. Influence of connected and autonomous vehicles on traffic flow stability and throughput. Transportation Research Part C: Emerging Technologies, 71, 143–163.
Talebpour, A., Mahmassani, H.S., 2014. Modeling acceleration behavior in a connected environment, in: Celebrating 50 Years of Traffic Flow Theory: A Symposium. Transportation Research Circular. Transportation Research Board, Portland, OR.
Talebpour, A., Mahmassani, H.S., Bustamante, F., E., 2016. Modeling Driver Behavior in a Connected Environment : Integrated Microscopic Simulation of Traffic and Mobile Wireless Telecommunication Systems. Transportation Research Record: Journal of the Transportation Research Board. Volume 2560, 75-86.
Talebpour, A., Mahmassani, H.S., Elfar, A., 2017. Investigating the Effects of Reserved Lanes for Autonomous Vehicles on Congestion and Travel Time Reliability. The 96th Transportation Research Board Annual Meeting. Washington D.C.
Treiber, M., Kesting, A., 2013. Traffic flow dynamics. Springer.
Treiber, M., Kesting, A., Helbing, D., 2007. Influence of reaction times and anticipation on stability of vehicular traffic flow. Transp. Res. Rec. 1999, 23–29. doi:10.3141/1999-03
van Arem, B., van Driel, C.J.G., Visser, R., Arem, B. Van, Driel, C.J.G. Van, 2006. The impact of cooperative adaptive cruise control on traffic-flow characteristics. Intell. Transp. Syst. IEEE Trans. 7, 429–436. doi:10.1109/TITS.2006.884615
Williams, K., 2013. Transportation planning consideration for automated vehicles.
Zhao, L., Sun, J., 2013. Simulation framework for vehicle platooning and car-following behaviors under connected-vehicle environment. Procedia - Soc. Behav. Sci. 96, 914–924. doi:http://dx.doi.org/10.1016/j.sbspro.2013.08.105
Zhou, F., Li, X., Ma, J., 2015. Parsimonious shooting heuristic for trajectory control of connected automated traffic part i: theoretical analysis with generalized time geography. arXiv:1511.04810 [math.OC].

Appendix A

Here we will prove that if a set of (α_1, α_2,…, α_K) maximizes total flow Q_K, all lanes should have reached their physical lane capacity; i.e.,

$$q_i = C_i = \frac{1}{1-\alpha_i\varepsilon_i}, \forall \ i = 1,2,\cdots K, K > 1 \ . \tag{A1}$$

For the proof, we consider three cases below, and we will show that none of these three cases is the optimal and thus the optimal solution form has to be in the format stated in (A1).

Case (i): in the solution set (q_1, q_2,…, q_K), there exists at least one lane that is empty.

Case (ii): in the solution set (q_1, q_2,…, q_K), there exists one and only one lane that is partially filled; i.e., it has not reached its physical lane capacity; and all other lanes are fully filled.

Case (iii): in the solution set $(q_1, q_2, \ldots, q_K)$, there exists two or more lanes that are partially filled, and all other lanes are fully filled.

Analysis of Case (i)

Let lane I be one of the empty lanes; i.e., $q_I = 0$. From flow conservation, we have

$$(p-\alpha_I)q_I + \sum_{\substack{i=1\\ i\neq I}}^{K}(p-\alpha_i)q_i = 0. \tag{A2}$$

In this case, the flow q_I does not matter and we can always increase q_I to C_I with the conservation valid. Thus, we can set $\alpha_I = p$ and increase q_I until $q_I = C_I$. This will increase the overall flow Q_K. Apply this process to all empty lanes. Towards the end, we will have all empty lanes become fully filled; i.e., in the format of (A1).

Analysis of Case (ii)

Let lane I be the partially filled lane and lane J be one of the lanes that's fully filled; i.e.,

$$0 < q_I < 1, q_J = C_J, I \neq J\ .$$

From flow conservation, we have

$$(p-\alpha_I)q_I + (p-\alpha_J)C_J + \sum_{\substack{i=1\\ i\neq I\\ i\neq J}}^{K}(p-\alpha_i)q_i = 0.$$

This can be reformulated as

$$(p-\alpha_I)q_I + (p-\alpha_J)C_J = €_2, \tag{A3}$$

where

$$€_2 = -\sum_{\substack{i=1\\ i\neq I\\ i\neq J}}^{K}(p-\alpha_i)q_i$$

We fix the flow and AV proportion of all other lanes except for I and J, then the RHS of Eqn. (A3), $€_2$, is a constant. If $\alpha_I = p$, we can increase q_I until $q_I = C_I$ with the flow conservation preserved, which will increase the total flow Q_K. The new set of $(q_1, q_2, \ldots, q_K)$ is consistent with Eqn. (A1). If $\alpha_I \neq p$, we consider changing α_J, and thus α_I and/or q_I as long as the flow conservation is valid; i.e., $€_2$ remains a constant. Let Q_{IJ} be the total flow of lane I and J; i.e., $Q_{IJ} = q_I + q_J$. We reformulate q_I and Q_{IJ} as below:

$$q_I = \frac{€_2-(p-\alpha_J)C_J}{p-\alpha_I}, \tag{A4a}$$

$$Q_{IJ} = q_I + C_J = \frac{-\alpha_I C_0 + €_2 + \alpha_J(C_0 - €_2\varepsilon_J)}{(-1+\alpha_J\varepsilon_J)(\alpha_I - p)}. \tag{A4b}$$

If we fix all other variables except for α_J, we can take the derivative of q_I, and Q_{IJ} in respect to α_J:

$$\frac{dq_I}{d\alpha_J} = C_0\frac{(1-\varepsilon_J p)}{(1-\alpha_J\varepsilon_J)^2(p-\alpha_I)}, \tag{A5a}$$

$$\frac{dQ_{IJ}}{d\alpha_J} = C_0\frac{(1-\alpha_I\varepsilon_J)}{(1-\alpha_J\varepsilon_J)^2(p-\alpha_I)}\ . \tag{A5b}$$

Clearly, if $p < \alpha_I \leq 1, \frac{dq_I}{d\alpha_J} < 0$ and $\frac{dQ_{IJ}}{d\alpha_J} < 0$ suggesting that q_I and also Q_{IJ} increase as α_J decreases. In this case, we will decrease α_J to increase q_I and Q_{IJ} until either (a) q_I equals to C_I or (b) $\alpha_J = 0$, whichever occurs first, during which the flow conservation always holds. The former case is consistent with Eqn. (A1). For the latter case, assuming that now the flow on lane I is $\bar{q}_I$, according to the flow conservation, we have

$$(p-\alpha_I)\bar{q}_I = €_2 - pC_0\ , \tag{A6}$$

where $p < \alpha_I \leq 1$, and $\bar{q}_I < C_I$.

Note that Eqn. (A6) suggests that the RHS should be a negative constant. Then we have

$$\bar{q}_I = \frac{€_2 - pC_0}{p - \alpha_I}, \tag{A7}$$

This suggests that $\bar{q}_I$ decreases with α_I. Thus, we can decrease α_I to increase $\bar{q}_I$. Note that eventually we can achieve $\bar{q}_I = C_I = 1/(1 - \alpha_I \varepsilon_I)$, because as α_I approaches p, the denominator is really small while the numerator remains constant, which can result in an infinite value on the RHS. Thus, the RHS is assure to reach C_I before α_I decreases to p. Therefore, this case now becomes the format of Eqn. (A1).

In the case that $0 \le \alpha_I < p$, the process is the opposite but in a similar manner. Specifically, we have $\frac{dq_I}{d\alpha_J} > 0$ and $\frac{dQ_{IJ}}{d\alpha_J} > 0$ suggesting that q_I and also Q_{IJ} increase as α_J increases. In this case, we will increase α_J to increase q_I and Q_{IJ} until either (a) q_I equals to C_I or (b) $\alpha_J = 1$, whichever occurs first. The former case is consistent with Eqn. (A1). For the latter case, the flow conservation form (originally in Eqn. (A6)) becomes

$$(p - \alpha_I)\bar{q}_I = €_2 - C_0 \, \frac{p-1}{1-\varepsilon_J}, \tag{A8}$$

Eqn. (A7) suggests that the RHS should be a positive constant. Then we have

$$\bar{q}_I = \frac{€_2 - C_0 \frac{p-1}{1-\varepsilon_J}}{p - \alpha_I} \tag{A9}$$

This suggests that $\bar{q}_I$ increases with α_I. Thus, we can increase α_I to increase $\bar{q}_I$ until $\bar{q}_I = C_I$, which is guaranteed due to the same reason explained above.
Thus, overall, a solution in Case (ii) can always be improved to increase Q_K and becomes the format of Eqn. (A1).

<u>*Analysis of Case (iii)*</u>

If any of these partially filled lanes, lane i, has AV proportion equal to p, according to the flow conservation, the flow q_i does not matter and we can always increase q_i to C_i. Therefore, we only need to consider the scenario where the partially filled lanes have AV proportion not equal to p. Let I and J ($i \neq j$) be two such lanes; i.e., $\alpha_I \neq p$ and $\alpha_J \neq p$. Then we have

$$0 < q_I < C_I \ ,$$
$$0 < q_J < C_J \ .$$

From flow conservation we

$$(p - \alpha_I)q_I + (p - \alpha_J)q_J + \sum_{\substack{i=1 \\ i \neq I \\ i \neq J}}^{K} (p - \alpha_i)q_i = 0$$

This can be reformulated as

$$(p - \alpha_I)q_I + (p - \alpha_J)q_J = €_3, \tag{A10}$$

where

$$€_3 = -\sum_{\substack{i=1 \\ i \neq I \\ i \neq J}}^{K} (p - \alpha_i)q_i \ .$$

If we fix the AV proportion and flow of all lanes except for I and J, then $€_3$ is a constant. Now consider changing the flow on lane I and J. Reformulating Eqn. (A10), we have

$$q_I = \frac{€_3 - (p - \alpha_J)q_J}{p - \alpha_I} \ .$$

$$Q_{IJ} = q_I + q_J = -\frac{€_3 - \alpha_I q_J + \alpha_J q_J}{\alpha_I - p} \ .$$

If we fix all other variables on the RHS except for q_J, we can get the derivatives of q_I and Q_{IJ} in respect to q_J, which are given as below:

$$\frac{dq_I}{dq_J} = -\frac{\alpha_J - p}{\alpha_I - p}$$

$$\frac{dQ_{IJ}}{dq_J} = 1 - \frac{\alpha_J - p}{\alpha_I - p}$$

These indicate that if $0 < \frac{\alpha_J - p}{\alpha_I - p} < 1$, q_I will decrease with q_J but Q_{IJ} will increase or remains constant (the case that $\frac{\alpha_J - p}{\alpha_I - p} = 1$) . In this case, we will increase q_J until either (a) $q_J = C_J$ or (b) $q_I = 0$, whichever occurs first. If (a) occurs first, the number of partially filled lanes now reduces by one and we can repeat this process until there is only one partially filled lane, which now becomes the Case (ii). If (b) occurs first, it becomes Case (i). With either case, eventually, the solution can be improved to increase Q_K and $(q_1, q_2, \ldots, q_K)$ will be in the format of Eqn. (A1).

If $\frac{\alpha_J - p}{\alpha_I - p} < 0$, q_I and Q_{IJ} will increase with q_J. In this case, we can increase q_J until either (a) $q_J = C_J$ or (b) $q_I = C_I$, whichever occurs first. In any case, the number of partially filled lanes now reduces by one. We can repeat this process until there is only one partially filled lane, which now becomes the Case (ii).

If $\frac{\alpha_J - p}{\alpha_I - p} > 1$, q_I and Q_{IJ} will decrease with q_J. In this case, we can decrease q_J to increase q_I and Q_{IJ} until either (a) $q_I = C_I$ or (b) $q_J = 0$, whichever occurs first. Similar to the scenario of $0 < \frac{\alpha_J - p}{\alpha_I - p} < 1$, if (a) occurs first, the number of partially filled lanes now reduces by one and we can repeat this process until there is only one partially filled lane, which now becomes the Case (ii). If (b) occurs first, it becomes Case (i).

In any case, the number of partially filled lanes now reduces by one. We can repeat this process until there is only one partially filled lane, which now becomes the Case (ii).

Therefore, it is clear that none of the three cases will maximize Q_K and the optimal solution has to be in the format of Eqn. (A1).

Appendix B

To prove that to achieve higher capacity, AVs should use the most efficient lane to the possible extent and then the less efficient lanes.

Proof:

According to the result of Appendix A, if the total flow is maximized, all lanes should reach their respective physical capacity. Let $\tilde{Q}$ denote the total flow when all lanes reach their respective physical lane capacities. Namely,

$$\begin{cases}\tilde{Q} = \sum_{i=1}^{K} C_i \\ p\tilde{Q} = \sum_{i=1}^{K} C_i \alpha_i\end{cases} \tag{B1}$$

where

$$C_i = \frac{1}{1-\alpha_i \varepsilon_i}, \forall \ i = 1,2, \cdots K, K > 1\,, \tag{B2}$$

$$0 \le \alpha_i \le 1, 0 \le p \le 1, \forall\, i = 1,2, \cdots, K \tag{B3}$$

$$\varepsilon_1 \ge \varepsilon_2 \ge \varepsilon_3 \ge \cdots \ge \varepsilon_K, 0 \le \varepsilon_i < 1, \ \forall\, i = 1,2, \cdots, K \tag{B4}$$

Assuming that the AV proportion on all lanes are fixed except for $i, j \in [1, K]$ where $i < j$.

Consider the capacity sum of lane i, j:

$$\delta = \frac{1}{1-\alpha_i \varepsilon_i} C_0 + \frac{1}{1-\alpha_j \varepsilon_j} C_0. \tag{B5}$$

Since the AV proportions on all other lanes (except i and j) are fixed, from (B1) it's clear that $\tilde{Q}$ depends on δ and when δ is maximized, $\tilde{Q}$ is maximized too.

According to the flow conservation in (B1), we have

$$\frac{p-\alpha_i}{1-\alpha_i \varepsilon_i} + \frac{p-\alpha_j}{1-\alpha_j \varepsilon_j} = \psi, \tag{B6}$$

where

$$\psi = -\sum_{t=1, t\neq i,, t\neq j}^{K} \frac{p-\alpha_t}{1-\alpha_t \varepsilon_t}$$

Notice that the term ψ is a constant because all AV proportions on all other lanes (except i and j) are fixed. Take the full derivative of Eqn. (B6), we have

$$\frac{\alpha_i'(\varepsilon_i p - 1)}{(1-\alpha_i \varepsilon_i)^2} + \frac{\alpha_j'(\varepsilon_j p - 1)}{(1-\alpha_j \varepsilon_i)^2} = 0. \tag{B7}$$

Similarly, we take the derivative of δ in (B5), we have

$$\delta' = C_0 \frac{\alpha_i' \varepsilon_i}{(1-\alpha_i \varepsilon_i)^2} + C_0 \frac{\alpha_j' \varepsilon_j}{(1-\alpha_j \varepsilon_i)^2}. \tag{B8}$$

Integrate Eqn. (B7-B8), we have

$$\delta' = -\frac{\varepsilon_i - \varepsilon_j}{(1-\varepsilon_i p)(1-\alpha_j \varepsilon_j)^2} C_0 \alpha_j'. \tag{A9}$$

Note that since if $i < j$, we have $\varepsilon_i \ge \varepsilon_j$. Also, since $0 \le \varepsilon_i, \varepsilon_j < 1$, $0 < \alpha_i, \alpha_j \le 1$ and $0 \le p \le 1$, the denominator is always positive.

Clearly, if $\varepsilon_i > \varepsilon_j$, to obtain an increasing trend of δ (i.e., $\delta' > 0$), we should decrease α_j and increase α_i (note that α_i' has an opposite sign of α_j' according to Eqn. (B7)). Considering the physical boundaries of α_i and α_j in (A4a), δ will achieve the maximum when $\alpha_i = 1$. When $\varepsilon_i = \varepsilon_j$, δ' equals to 0, suggesting that δ doesn't change with α_i or α_j, which is expected because now the two lanes have the same efficiency. In this case, there exists a deterministic relationship between α_i and α_j that results in a contour of δ.

The results above imply that to maximize $\tilde{Q}$, we should assign AVs to the most efficient lane (with the largest ε) whenever possible.

Appendix C

Part I: in this part, we will show that the statement in Appendix A holds if α_i and the AV gain $E_i(\alpha_i)$ are correlated. Unless specified the value, we use E_i to refer to $E_i(\alpha_i)$.

Similar to the proof in Appendix, we consider the same three cases. Note that the proof for Case (i) and (iii) still hold when α_i and E_i are correlated. Thus, we only need to consider Case (ii).

<u>*Analysis of Case (ii)*</u>

Similar to the setting in Appendix A, let lane I be the partially filled lane and lane J be one of the lanes that's fully filled; i.e.,

$$0 < q_I < 1, q_J = C_J, I \neq J \,.$$

Note that (A3) and (A4) hold. We denote them using the new equation numbering:

$$(p - \alpha_I)q_I + (p - \alpha_J)C_J = 0, \tag{C1}$$

$$q_I = \frac{-(p-\alpha_J)}{p-\alpha_I} C_J \,, \tag{C2a}$$

$$Q_{IJ} = q_I + C_J = C_J \left(\frac{\alpha_J - \alpha_I}{p-\alpha_I}\right) = C_0 \frac{\alpha_J - \alpha_I}{(1-\alpha_J \varepsilon_J)(p-\alpha_I)} \,. \tag{C2b}$$

If we fix all other variables except for α_J, we can take the derivative of q_I, and Q_{IJ} in respect to α_J:

$$\frac{dq_I}{d\alpha_J} = C_0 \frac{1}{(1-\alpha_J E_J)(p-\alpha_I)} \left[1 + (\alpha_J - p)\frac{(\alpha_J E_J)'}{1-\alpha_J E_J}\right] \,, \tag{C3a}$$

$$\frac{dQ_{IJ}}{d\alpha_J} = C_0 \frac{1}{(1-\alpha_J E_J)(p-\alpha_I)} \left[1 + (\alpha_J - \alpha_I)\frac{(\alpha_J E_J)'}{1-\alpha_J E_J}\right] \,. \tag{C3b}$$

where $(\alpha_J E_J)'$ denotes the derivative of $\alpha_J E_J$ in respect to α_J, which is positive; i.e., $(\alpha_J E_J)' > 0$.

If $\alpha_J > p > \alpha_I$, $\frac{dq_I}{d\alpha_J} > 0$ and $\frac{dQ_{IJ}}{d\alpha_J} > 0$. We can increase α_J to increase q_I and also Q_{IJ} until (a) $q_I = C_I$ or (b) $\alpha_J = p$, whichever occurs first. Note that when (a) occurs, α_J is still larger than p according to (C2a). Thus, it infers that (a) will occur first, and thus the solution becomes the format in Eqn. (A1).

Consider the scenario that $\alpha_J < p < \alpha_I$. In this case, we have $(\alpha_J - p)\frac{(\alpha_J E_J)'}{1-\alpha_J E_J} > (\alpha_J - \alpha_I)\frac{(\alpha_J E_J)'}{1-\alpha_J E_J}$.

If $(\alpha_J - p)\frac{(\alpha_J E_J)'}{1-\alpha_J E_J} < -1$, we have $(\alpha_J - \alpha_I)\frac{(\alpha_J E_J)'}{1-\alpha_J E_J} < -1$ too, suggesting that $\frac{dq_I}{d\alpha_J} > 0$ and $\frac{dQ_{IJ}}{d\alpha_J} > 0$. In this case, similarly, we can increase α_J to increase q_I and also Q_{IJ} until (a) $q_I = C_I$ or (b) $\alpha_J = p$, whichever occurs first. As mentioned above, (a) will occur first.

If $(\alpha_J - \alpha_I)\frac{(\alpha_J E_J)'}{1-\alpha_J E_J} > -1$, we have $(\alpha_J - p)\frac{(\alpha_J E_J)'}{1-\alpha_J E_J} > -1$, suggesting that $\frac{dq_I}{d\alpha_J} < 0$ and $\frac{dQ_{IJ}}{d\alpha_J} < 0$. In this case, we can decrease α_J to increase q_I and also Q_{IJ} until (a) $q_I = C_I$ or (b) $\alpha_J = 0$, whichever occurs first. If (a) occurs first, the solution becomes the format of Eqn. (A1). If (b) occurs first, now we fix $\alpha_J = 0$. Then C_J is a constant. According to (C2a), q_I decreases with α_I. Given that, we can decrease α_I to increase q_I until (a) $q_I = C_I$ or (b) $\alpha_I = p$. For the (b) case, notice that when α_I approaches p, the RHS of (C2a) becomes a very large number. Therefore, it infers that (a) will occur first.

If $(\alpha_J - p)\frac{(\alpha_J E_J)'}{1-\alpha_J E_J} > -1$, and $(\alpha_J - \alpha_I)\frac{(\alpha_J E_J)'}{1-\alpha_J E_J} < -1$, we have that $\frac{dq_I}{d\alpha_J} < 0$ and $\frac{dQ_{IJ}}{d\alpha_J} > 0$. In this case, we can increase α_J to increase Q_{IJ} (but q_I will decrease) until (a) $q_I = 0$ or (b) $\alpha_J = p$. Note that (a) and (b) will occur simultaneously according to (C2a). This reduces to Case (i).

Thus, in Case (ii), the solution can be improved to increase Q_{IJ}, and thus the total flow.

Part II: in this part, we will perform a marginal analysis for $\tilde{Q}$ for a two-lane highway to obtain more insights.

As presented in Section 5, for the correlation case, we consider a two-lane highway. Assume that ε_1 (ε_2) is a function of α_1 (α_2); i.e., $\varepsilon_1 = E_1(\alpha_1)$ and $\varepsilon_2 = E_2(\alpha_2)$. Let $\tilde{Q}$ denote the total flow when both lanes reach their respective physical lane capacities. Then, we can formulate an optimization problem to maximize $\tilde{Q}$ (objective function) with respect to α_1 and α_2 (decision variables):

$$\tilde{Q} = C_0 \frac{1}{1-\alpha_1 E_1} + C_0 \frac{1}{1-\alpha_2 E_2}, \tag{D1}$$

subject to the flow conservation and physical boundaries of α_1, α_2, and p:

$$C_0 \frac{p-\alpha_1}{1-\alpha_1 E_1} + C_0 \frac{p-\alpha_2}{1-\alpha_2 E_2} = 0, \tag{D2}$$

$$0 \le \alpha_1 \le 1, \tag{D3a}$$

$$0 \le \alpha_2 \le 1, \tag{D3b}$$

$$0 \le p \le 1. \tag{D3c}$$

Note that this is not a convex optimization problem because the constraint for the flow conservation (Eqn. (D2)) is not necessarily convex (notice that E_1 and E_2 are functions of α_1 and α_2). Next we conduct a marginal analysis for $\tilde{Q}$.

Notice that from (D2), for a given p value and specified $E_1(\alpha_1)$, and $E_2(\alpha_2)$ functions, using numerical method, we can obtain a feasible solution domain of (α_1, α_2), denoted by π; i.e.,

$$\pi = \left\{(\alpha_1, \alpha_2), \middle| C_0 \frac{p-\alpha_1}{1-\alpha_1 E_1} + C_0 \frac{p-\alpha_2}{1-\alpha_2 E_2} = 0, 0 \le \alpha_1 \le 1, 0 \le \alpha_2 \le 1, 0 \le p \le 1\right\}. \tag{D4}$$

We take the derivative of the flow conservation (D2) in respect to α_2, which results in:

$$\frac{d\left(C_0 \frac{p-\alpha_1}{1-\alpha_1 E_1}\right)}{d\alpha_1} \frac{d\alpha_1}{d\alpha_2} = -\frac{d\left(C_0 \frac{p-\alpha_2}{1-\alpha_2 E_2}\right)}{d\alpha_2}, \tag{D5}$$

Clearly, α_1 is a function of α_2 but its format is too complex to have the explicit form. Let $A' = \frac{d\alpha_1}{d\alpha_2}$. We have

$$A' = -\frac{1-\alpha_1 E_1}{1-\alpha_2 E_2} * \left\{\left[1 + \frac{d(\alpha_2 E_2)/d\alpha_2}{1-\alpha_2 E_2}(\alpha_2 - p)\right] \Big/ \left[1 + \frac{d(\alpha_1 E_1)/d\alpha_1}{1-\alpha_1 E_1}(\alpha_1 - p)\right]\right\}. \tag{D6}$$

In the similar spirit, we take the derivative of C_1, C_2 and $\tilde{Q}$ in respect to α_2, and obtain the results below:

$$\frac{dC_1}{d\alpha_2} = \frac{dC_1}{d\alpha_1}\frac{d\alpha_1}{d\alpha_2} = \frac{d(\alpha_1 E_1)/d\alpha_1}{(1-\alpha_1 E_1)^2} A'. \tag{D7a}$$

$$\frac{dC_2}{d\alpha_2} = \frac{d(\alpha_2 E_2)/d\alpha_2}{(1-\alpha_2 E_2)^2}. \tag{D7b}$$

$$\frac{d\tilde{Q}}{d\alpha_2} = \frac{dC_1}{d\alpha_2} + \frac{dC_2}{d\alpha_2} = \frac{d(\alpha_1 E_1)/d\alpha_1}{(1-\alpha_1 E_1)^2} A' + \frac{d(\alpha_2 E_2)/d\alpha_2}{(1-\alpha_2 E_2)^2}. \tag{D8}$$

Note that

$$d(\alpha_1 E_1)/d\alpha_1 = E_1 + \alpha_1 E_1', \tag{D9a}$$

$$d(\alpha_2 E_2)/d\alpha_2 = E_2 + \alpha_2 E_2', \tag{D9b}$$

where E_1' (E_2') is the derivative of E_1 (E_2) in respect to α_1 (α_2).

From the perspective of meaningful physical interpretation, here we assume that $E_1' \ge 0$, $E_2' \ge 0$, and $d(\alpha_2 E_2)/d\alpha_2 > 0$.

Next, we aim to see the possible values of $\frac{d\tilde{Q}}{d\alpha_2}$. From flow conservation (D2), there are three possible relations between α_1 and α_2:

Case (1): $\alpha_1 = p = \alpha_2$

Case (2): $\alpha_1 > p > \alpha_2$

Case (3): $\alpha_2 > p > \alpha_1$

In Case (1), $\tilde{Q}$ is determined, given below:

$$\tilde{Q} = C_0 \frac{1}{1-pE_1(p)} + C_0 \frac{1}{1-pE_2(p)}. \tag{D10}$$

Next, we analyze Case (2) and (3).

Analysis for Case (2)

There are three subcases illustrated below.

Case (2-1): if $A' > 0$, from (D7-D9) we have $\frac{dC_1}{d\alpha_2} > 0, \frac{dC_2}{d\alpha_2} > 0$, and $\frac{d\tilde{Q}}{d\alpha_2} > 0$. In this case, we should increase α_2 to increase C_1, C_2, and $\tilde{Q}$ in π (the feasible solution domain of (α_1, α_2). Namely, we should increase α_2 to the maximum value in the feasible domain defined by the following conditions:

$$\begin{cases} \frac{p-\alpha_1}{1-\alpha_1 E_1} + \frac{p-\alpha_2}{1-\alpha_2 E_2} = 0 \\ 1 \geq \alpha_1 > p > \alpha_2 \geq 0 \\ A' > 0 \end{cases}, \tag{D11}$$

where A' is given in (D6). This domain is equivalent to the following after plugging (D6):

$$\begin{cases} \frac{p-\alpha_1}{1-\alpha_1 E_1} + \frac{p-\alpha_2}{1-\alpha_2 E_2} = 0 \\ 1 \geq \alpha_1 > p > \alpha_2 \geq 0 \\ \frac{d(\alpha_2 E_2)/d\alpha_2}{1-\alpha_2 E_2} > \frac{1}{p-\alpha_2} \end{cases}. \tag{D12}$$

If $A' < 0$, from (D7-D9) we have $\frac{dC_1}{d\alpha_2} < 0, \frac{dC_2}{d\alpha_2} > 0$, but it is uncertain whether $\frac{d\tilde{Q}}{d\alpha_2}$ is larger or smaller than 0.

Case (2-2): $A' < 0$, and $\frac{d\tilde{Q}}{d\alpha_2} > 0$. In this case, to increase $\tilde{Q}$ we should increase α_2 to the maximum value in the feasible domain defined by the following conditions:

$$\begin{cases} \frac{p-\alpha_1}{1-\alpha_1 E_1} + \frac{p-\alpha_2}{1-\alpha_2 E_2} = 0 \\ 1 \geq \alpha_1 > p > \alpha_2 \geq 0 \\ A' < 0 \\ \frac{d\tilde{Q}}{d\alpha_2} > 0 \end{cases}, \tag{D13}$$

where A' is given in (D6) and $\frac{d\tilde{Q}}{d\alpha_2}$ is given in (D8). This domain is equivalent to the following after plugging (D6) and (D8):

$$\begin{cases} \frac{p-\alpha_1}{1-\alpha_1 E_1} + \frac{p-\alpha_2}{1-\alpha_2 E_2} = 0 \\ 1 \geq \alpha_1 > p > \alpha_2 \geq 0 \\ p < \frac{1-\alpha_2 E_2}{d(\alpha_2 E_2)/d\alpha_2} + \alpha_2 < \frac{1-\alpha_1 E_1}{d(\alpha_1 E_1)/d\alpha_1} + \alpha_1 \end{cases}. \tag{D14}$$

Case (2-3): $A' < 0$, and $\frac{d\tilde{Q}}{d\alpha_2} < 0$. In this case, to increase $\tilde{Q}$ we should decrease α_2 to the minimum value in the feasible domain defined by the following conditions:

$$\begin{cases} \frac{p-\alpha_1}{1-\alpha_1 E_1} + \frac{p-\alpha_2}{1-\alpha_2 E_2} = 0 \\ 1 \geq \alpha_1 > p > \alpha_2 \geq 0 \\ A' < 0 \\ \frac{d\tilde{Q}}{d\alpha_2} < 0 \end{cases}, \tag{D15}$$

where A' is given in (D6) and $\frac{d\tilde{Q}}{d\alpha_2}$ is given in (D8). This domain is equivalent to the following after plugging (D6) and (D8):

$$\begin{cases} \frac{p-\alpha_1}{1-\alpha_1 E_1} + \frac{p-\alpha_2}{1-\alpha_2 E_2} = 0 \\ 1 \geq \alpha_1 > p > \alpha_2 \geq 0 \\ \frac{1-\alpha_2 E_2}{d(\alpha_2 E_2)/d\alpha_2} + \alpha_2 > \frac{1-\alpha_1 E_1}{d(\alpha_1 E_1)/d\alpha_1} + \alpha_1 \end{cases}. \tag{D16}$$

The feasible domains for the three subcases can be obtained numerically, and with that, the optimal solution of (α_1, α_2) to the maximum $\tilde{Q}$ can be obtained numerically too.

Analysis for Case (3)

The analysis for Case (3) is symmetric to Case (2) and the feasible domains and optimal solutions can be obtained in the similar way.

Note that with the optimal solutions from the three cases, we need to compare their maximum $\tilde{Q}$ to select the final optimal solution of (α_1, α_2).

One can see that the feasible domains and the optimal solutions depend on the $E_i(\alpha_i)$ functions and they are usually too complex to have the explicit analytical form, even assuming the simplest linear form of $E_i(\alpha_i)$.

(a)

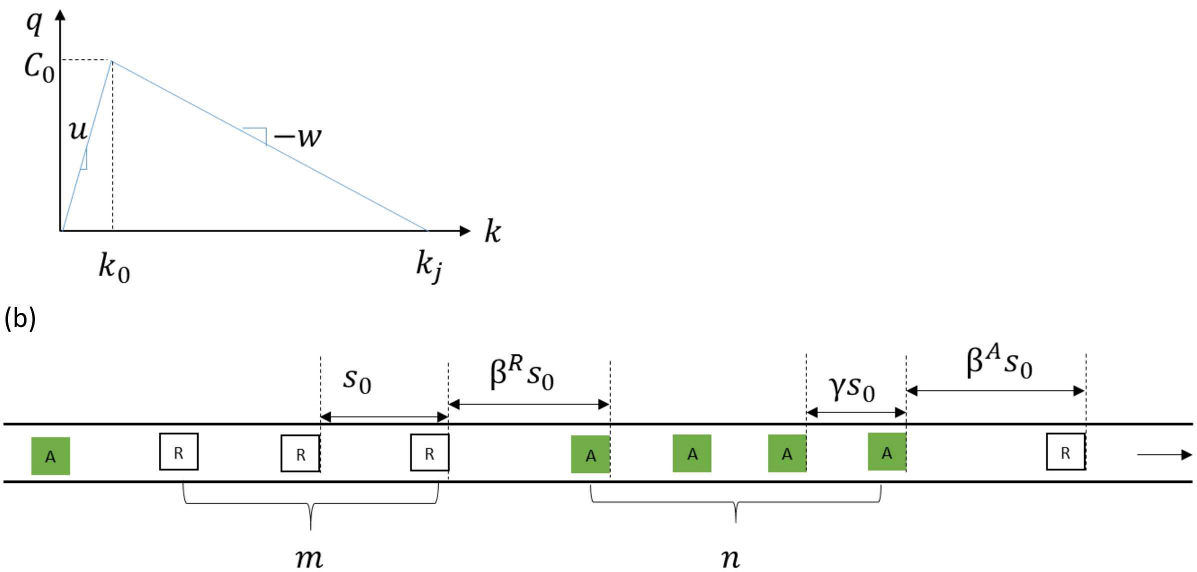


Fig. 2-1: (a) Fundamental diagram; (b) illustration of inter-vehicle spacing characteristics

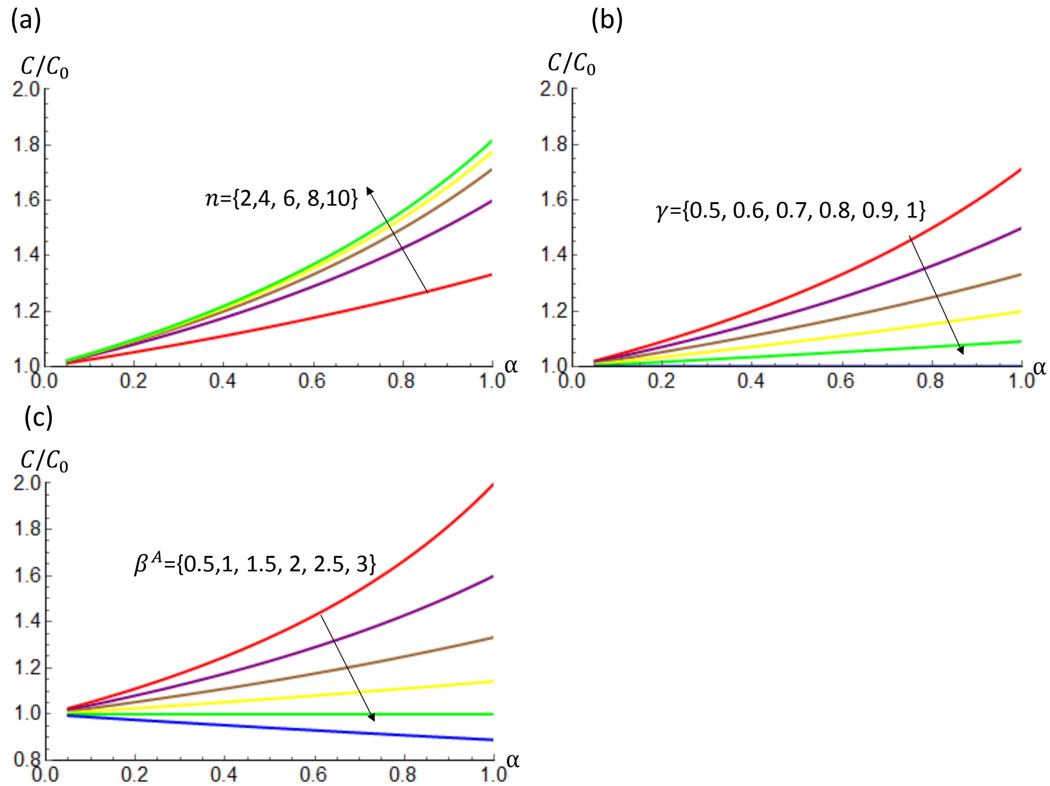


Fig. 2-2: (a) Capacity change with various n ($\gamma = 0.5, \beta^A = \beta^R = 1$); (b) Capacity change with various γ ($n = 6, \beta^A = \beta^R = 1$); (c) Capacity change with various β^A ($n = 4, \gamma = 0.5, \beta^R = 1$).

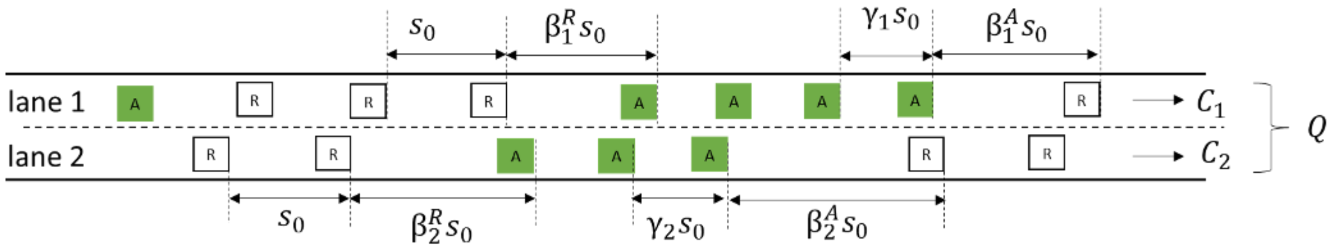


Fig. 3-1: Sketch of two-lane framework.

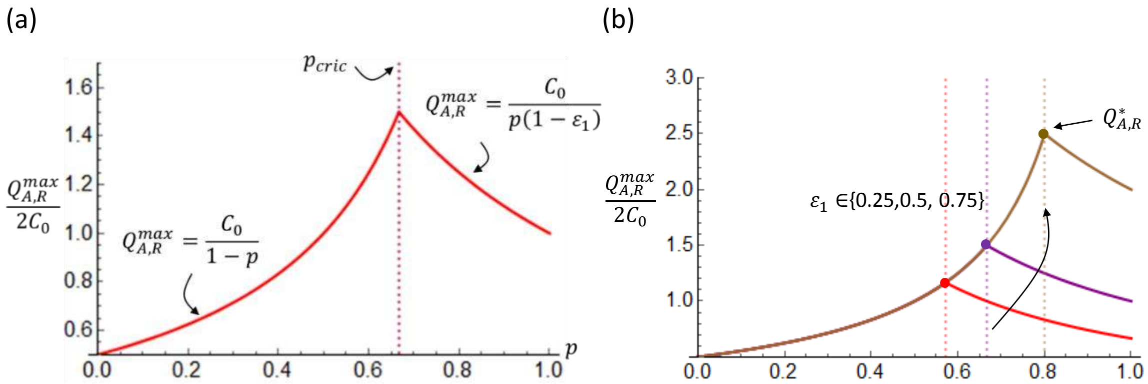


Fig. 3-2: (A, R) policy: (a) $Q_{A,R}^{max}$ under various p ($\varepsilon_1 = 0.5, \varepsilon_2 = 0.2$); (b) impacts of ε_1 on $Q_{A,R}^{max}$ and $Q_{A,R}^{*}$ ($\varepsilon_2 = 0.2$).

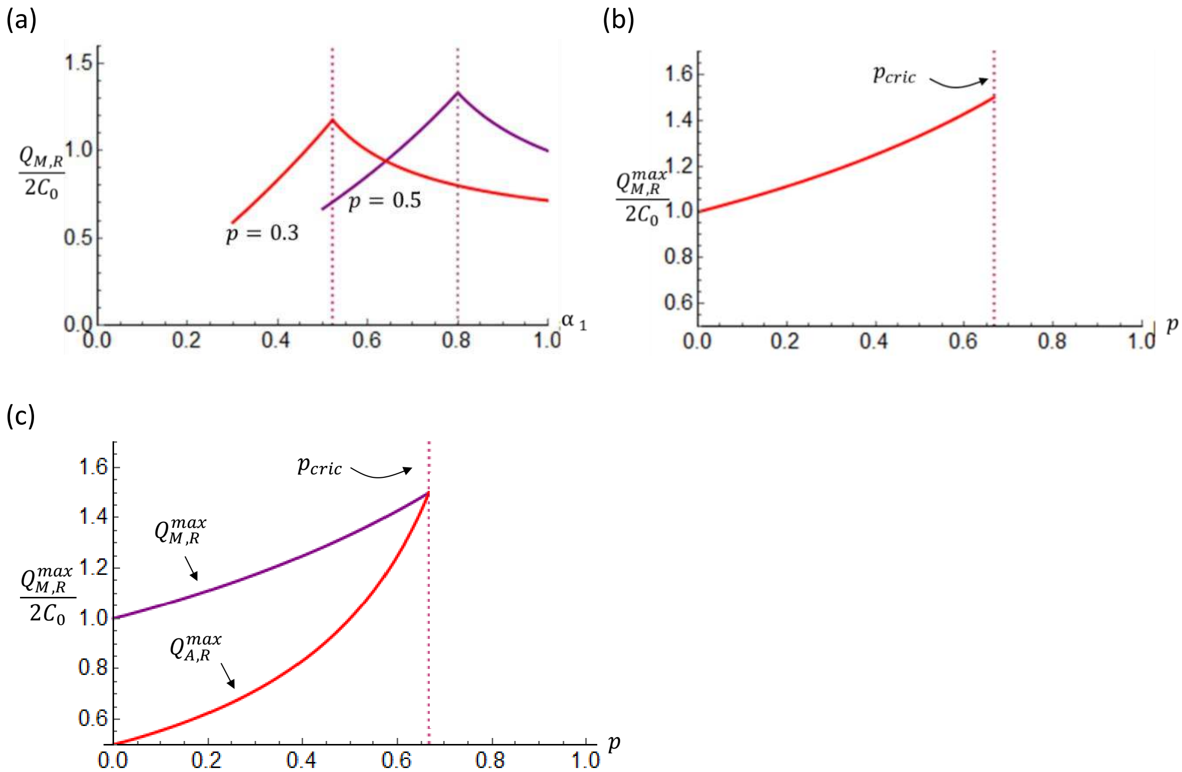
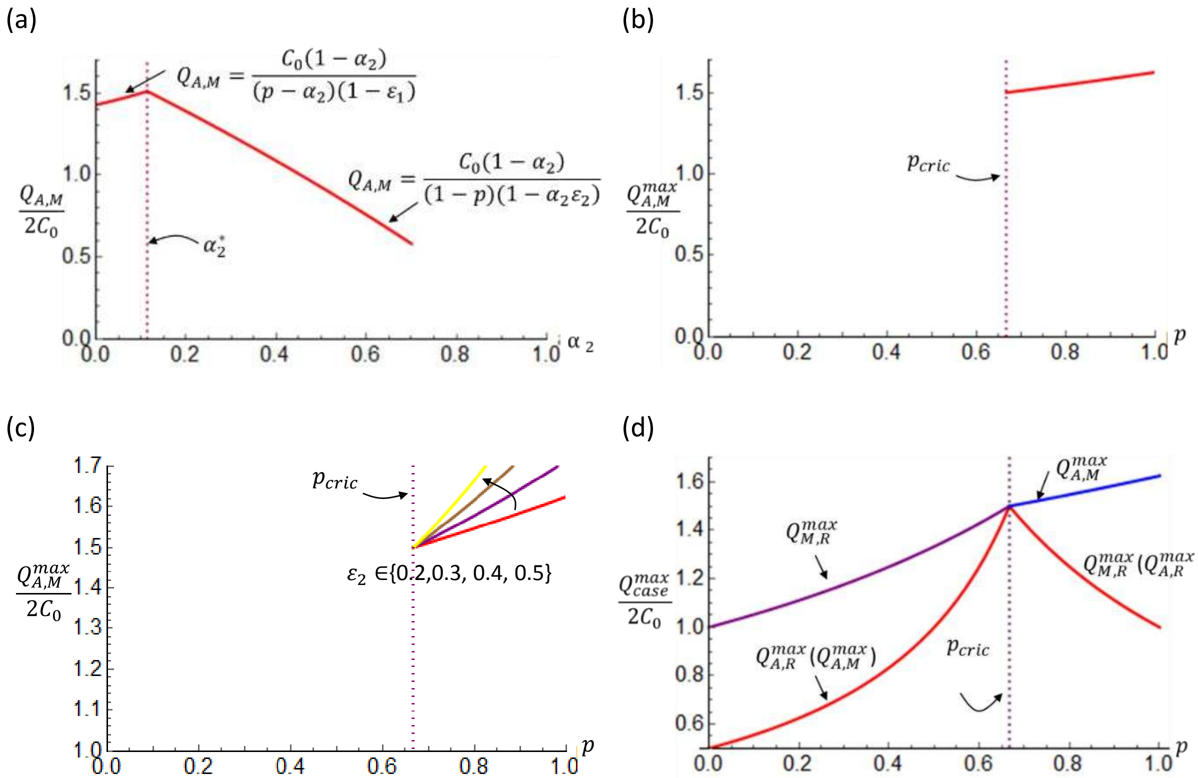
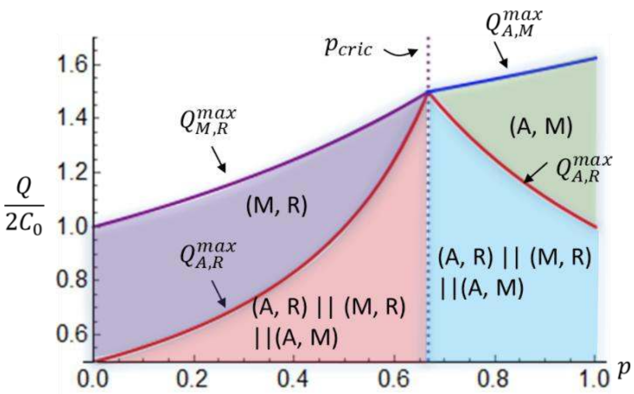


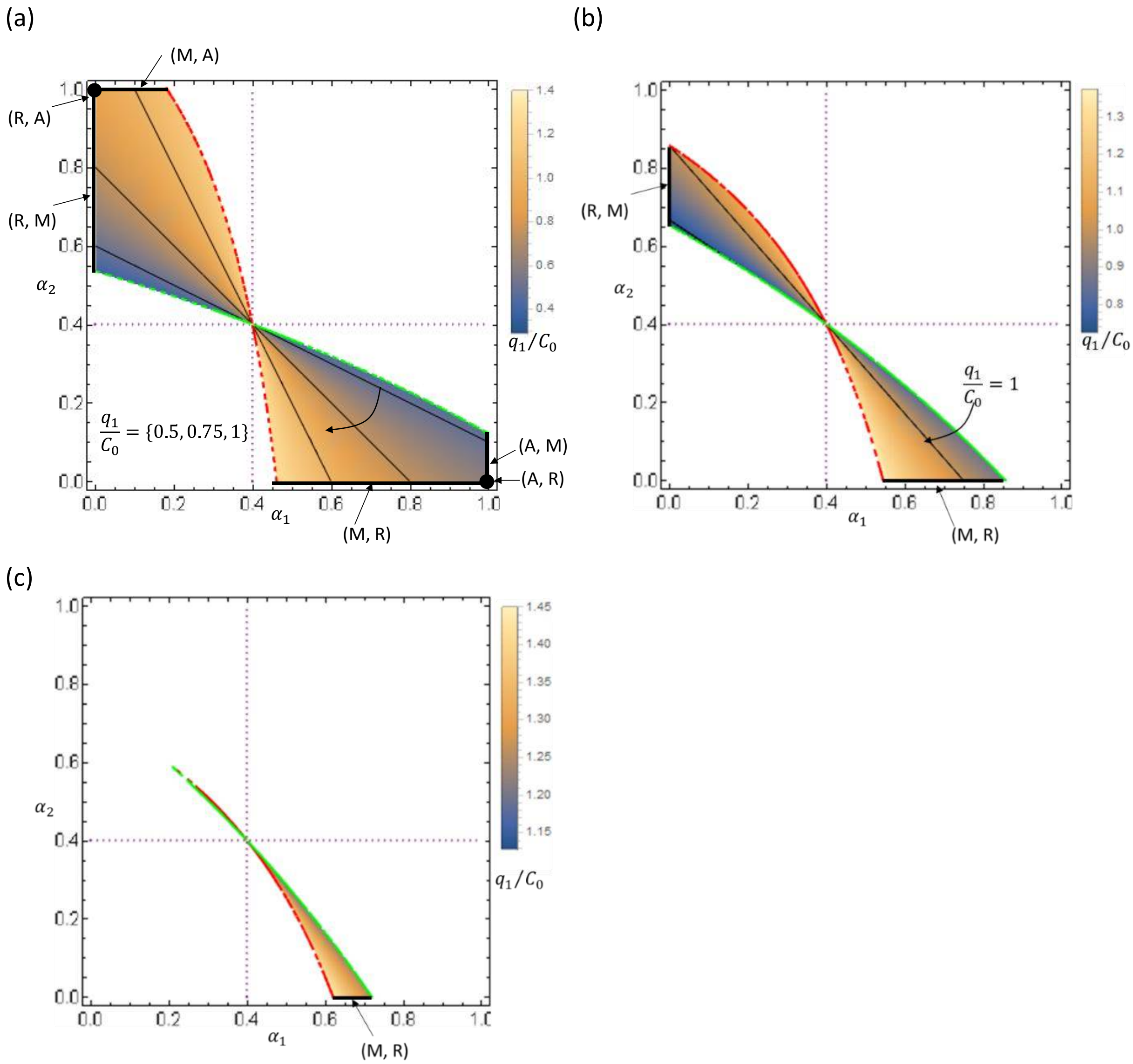
Fig. 3-3: (M, R) policy ($\varepsilon_1 = 0.5, \varepsilon_2 = 0.2$): (a) $Q_{M,R}$ under various α_1 (vertical line denotes $\alpha_1 = \frac{2p}{1+p\varepsilon_1}$) ; (b) $Q_{A,R}^{max}$ under various p; (c) comparison of $Q_{A,R}^{max}$ and $Q_{M,R}^{max}$.



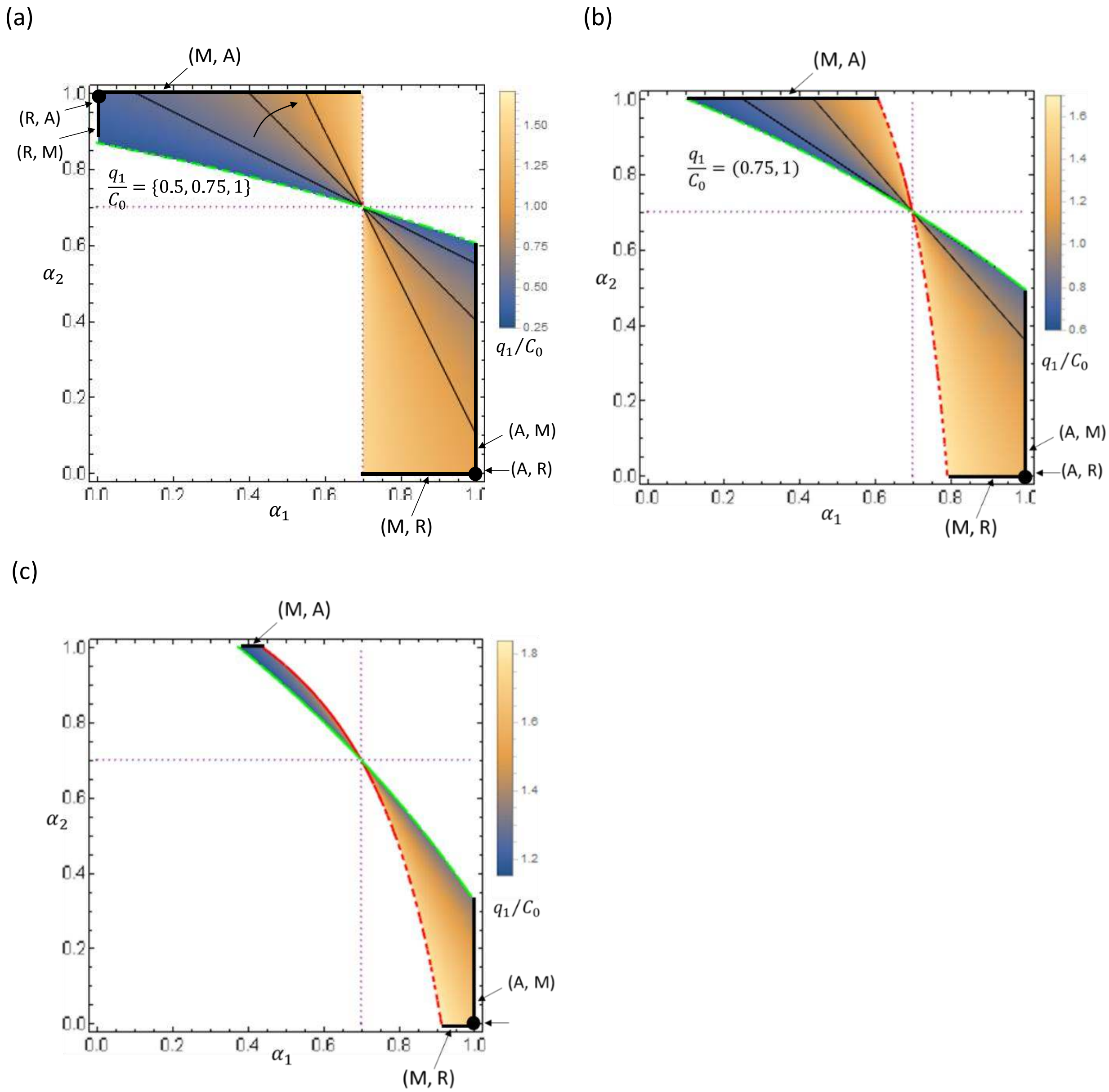
10 Fig. 3-4: (A,M) policy ($\varepsilon_1 = 0.5, \varepsilon_2 = 0.2$): (a) $Q_{A,M}$ under various α_2; (b) $Q_{A,M}^{max}$ under various p; (c) $Q_{A,M}^{max}$ under various ε_2;(c) comparison of $Q_{A,R}^{max}$, $Q_{M,R}^{max}$, and $Q_{A,M}^{max}$.



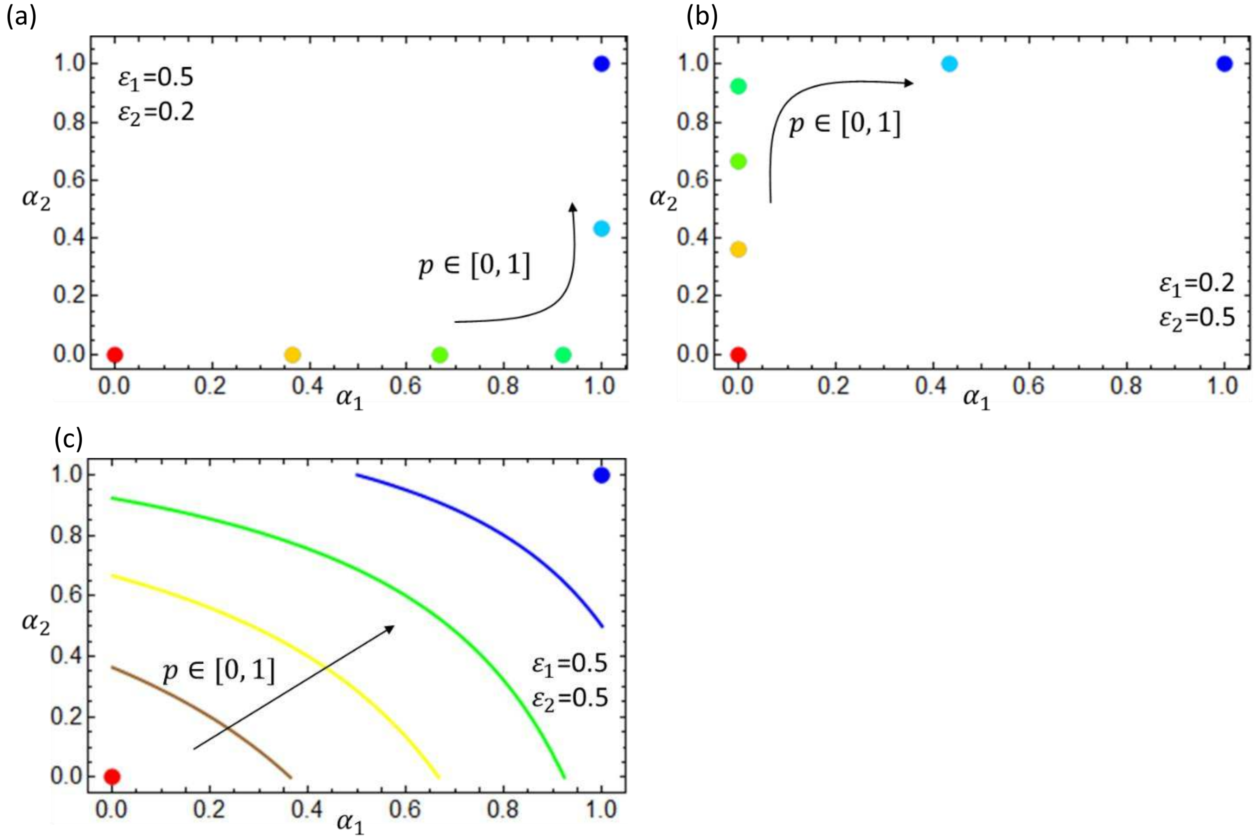
11 Fig. 3-5: Feasible policies under various demand.



12 Fig. 3-6: Valid domain for $p < p_{cric}$ ($p = 0.4, \varepsilon_1 = 0.5, \varepsilon_2 = 0.2$): (a) low demand ($\frac{Q}{2C_0} = 0.75$); (a) medium demand ($\frac{Q}{2C_0} = 0.94$); (c) high demand ($\frac{Q}{2C_0} = 1.125$).



13 Fig. 3-7: Valid domain for $p > p_{cric}$ ($p = 0.7, \varepsilon_1 = 0.5, \varepsilon_2 = 0.2$): (a) low demand ($\frac{Q}{2C_0} = 0.75$); (a) medium demand ($\frac{Q}{2C_0} = 0.94$); (c) high demand ($\frac{Q}{2C_0} = 1.125$).



14 Fig. 3-8: Solution to capacity under various p: (a) scenario of $\varepsilon_1 > \varepsilon_2$; (a) scenario of $\varepsilon_1 < \varepsilon_2$; (c) scenario of $\varepsilon_1 = \varepsilon_2$.

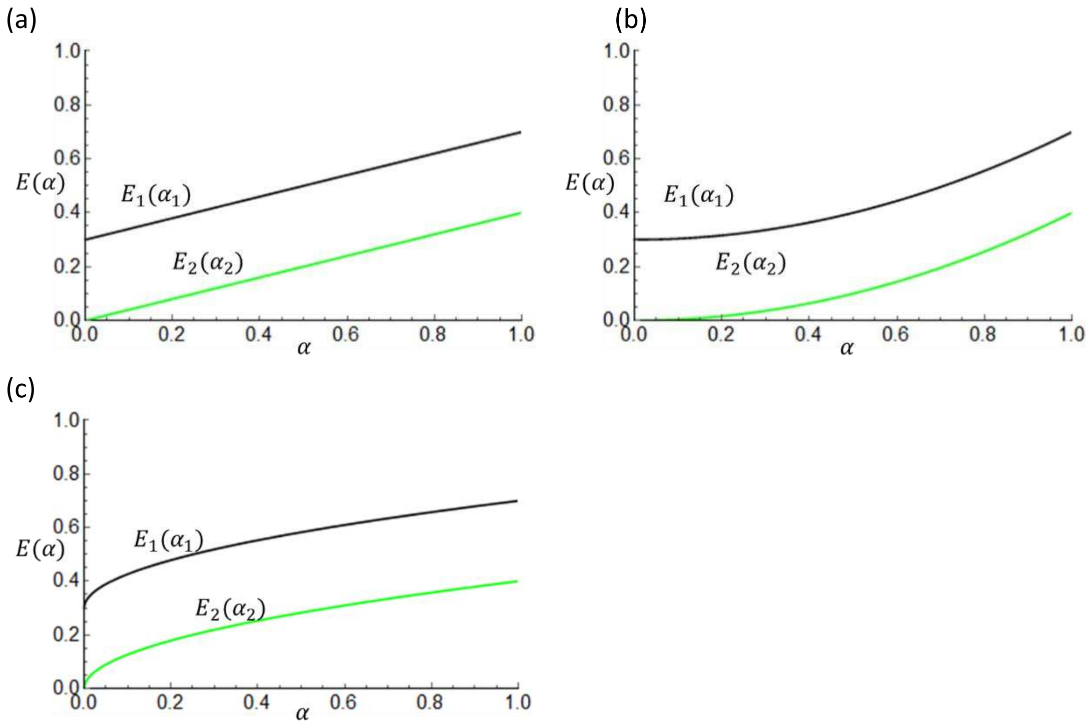


Fig. 5-1: Four functions for $E_i(\alpha_i)$: (a) linear function ($E_1(\alpha_1) = 0.3 + 0.4\alpha_1, E_2(\alpha_2) = 0.4\alpha_2$); (b) convex function ($E_1(\alpha_1) = 0.3 + 0.4{\alpha_1}^2, E_2(\alpha_2) = 0.4{\alpha_2}^2$); (c) concave function ($E_1(\alpha_1) = 0.3 + 0.4\sqrt{\alpha_1}, E_2(\alpha_2) = 0.4\sqrt{\alpha_2}$).

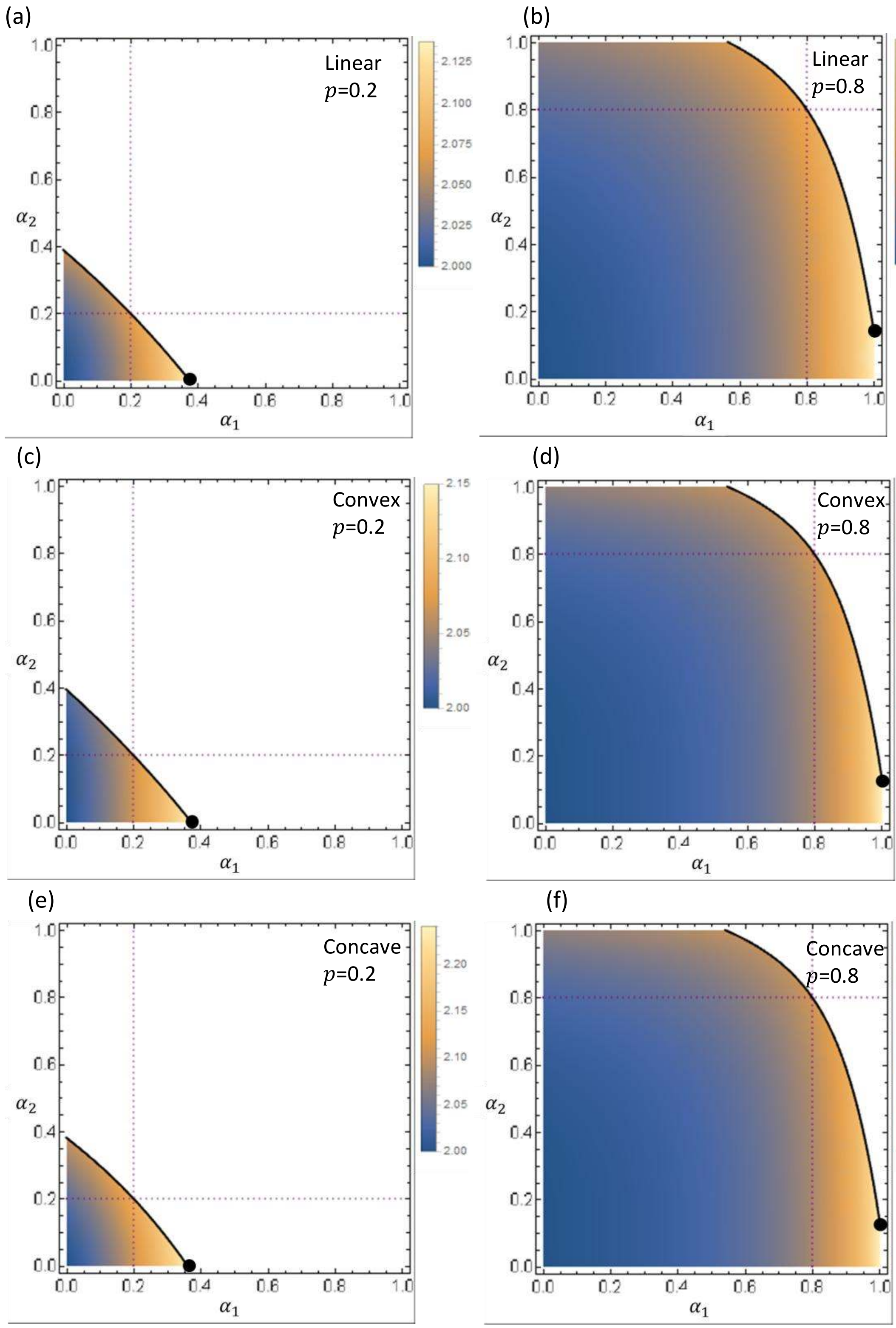


Fig. 5-2: Feasible domains and optimal solution with correlated α_i and $E_i(\alpha_i)$: (a-b) linear function ($E_1(\alpha_1) = 0.3 + 0.4\alpha_1$, $E_2(\alpha_2) = 0.4\alpha_2$); (c-d) convex function ($E_1(\alpha_1) = 0.3 + 0.4{\alpha_1}^2$, $E_2(\alpha_2) = 0.4{\alpha_2}^2$); (e-f) concave function ($E_1(\alpha_1) = 0.3 + 0.4\sqrt{\alpha_1}$, $E_2(\alpha_2) = 0.4\sqrt{\alpha_2}$); $p = 0.2$ for (a, c, e) and $p = 0.8$ for (b, d, f).