IEEE Open Journal of
Intelligent Transportation
Systems

Received 14 March 2024; revised 24 April 2024; accepted 4 May 2024. Date of publication 8 May 2024; date of current version 23 May 2024.

*Digital Object Identifier 10.1109/OJITS.2024.3397959*

# Dilemma of Responsibility-Sensitive Safety in Longitudinal Mixed Autonomous Vehicles Flow: A Human-Driver-Error-Tolerant Driving Strategy

**HONGSHENG QI**

College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310058, China

CORRESPONDING AUTHOR: H. QI (e-mail: qihongsheng@zju.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 52272314 and Grant 52131202; in part by The Ministry of Education in China Project of Humanities and Social Science under Grant 21YJCZH116; and in part by the Zhejiang Province Public Welfare Scientific Research Project under Grant LGF22E080007.

**ABSTRACT** The safety of autonomous vehicles (AVs) is a critical consideration for their widespread adoption. Responsibility sensitive safety (RSS) is proposed to serve as a model checking tool for AV safety. However, RSS alone cannot guarantee safety when they are mixed with human-driven vehicles (HDVs). These HDVs may disregard safety rules, creating dilemmas for AVs where they must choose between crashing into their leader or crashing into their follower. This manuscript defines this dilemma regarding the longitudinal driving and extends it to platooning scenarios with an arbitrary number of vehicles, referred to as polylemma. In polylemma, a violation of safety rules by one vehicle inevitably results in at least one crash between neighboring vehicles. To avoid the polylemma scenario, the manuscript proposes a human error-tolerant (HET) driving strategy, wherein AVs maintain an additional gap and prepare for moderate deceleration to account for potential errors by human drivers. The manuscript derives the risk reduction and capacity variation resulting from the implementation of this strategy at a given market penetration rate (MPR) using real world trajectory data. The analysis indicates that a 50% MPR would reduce risks due to human error by 80%, with a decrease in capacity which vary different for background traffic flow speed.

**INDEX TERMS** Mixed autonomous vehicles, responsibility sensitive safety, human driver error tolerant, polylemma avoidance, longitudinal driving strategy.

## I. INTRODUCTION

### A. BACKGROUND AND MOTIVATIONS

AUTONOMOUS vehicle technology has garnered significant interest from researchers and has been the focus of numerous studies [1], [2], [3], [4], [5]. Among the various advantages it offers [6], ensuring traffic safety is of utmost importance in order to gain consumer trust and adoption [7]. Multiple research efforts have addressed this issue [8], [9], and the concept of Responsibility Sensitive Safety has emerged as a formal framework to fulfill this purpose [10], [11], [12], [13], [14]. Recently, RSS has been further developed and applied to a wide range of scenarios [5], [15], [16], [17], [18], [19], [20].

Despite the intended purpose of RSS in ensuring the safety of autonomous vehicles, it is important to note that absolute prevention of collisions is unattainable [20], [21], [22], [23] due to the inherent limitations of directly controlling human-driven vehicles. HDVs may exhibit risky or erroneous behavior, as outlined by [3], [24], [25], and present considerable heterogeneity regarding risk levels [26]. Even if AVs strictly adhere to the rules specified by RSS or other safety guidelines, the presence of HDVs introduces the possibility of collisions. Surrounding HDVs may disregard safety regulations, forcing AVs into unavoidable collision scenarios. To mitigate this risk, one potential solution

The review of this article was arranged by Associate Editor Yajie Zou.

is to leverage AV movements to indirectly reduce the risk [27]. For example, the application of AVs in absorbing traffic waves in mixed traffic flows has garnered significant attention, as discussed by [28].

To illustrate the scenario where HDV error is introduced and our strategy, we consider a simple scenario involving a platoon of three vehicles, with the AV positioned in the middle. Suppose the last vehicle, an HDV, follows the AV too closely, disregarding the RSS rule. In such a situation, if the lead vehicle comes to a complete stop, the collision between the lead vehicle and the AV can be avoided because the AV adheres to the RSS rule. However, due to the smaller gap between the AV and its follower, a collision between the AV and the last vehicle becomes inevitable. To prevent a crash with its follower, the AV must drive with a moderate deceleration, thus leaving sufficient space for the last HDV to stop. However, this may result in a collision with the lead vehicle. We introduce the term "dilemma" to describe this particular circumstance. To avoid the dilemma, the AV needs to maintain an additional distance and decelerate at a moderate rate, as compared to the deceleration specified in RSS. By adjusting the AV's behavior in this manner, accommodation can be made for human driver errors, thereby reducing the risk associated with traffic flow operation. According to the National Highway Traffic Safety Administration (NHTSA), 94% of all traffic accidents are caused or influenced by human error [29], [30]. By proposing a driving strategy that tolerates human driver errors, a significant portion of the risks stemming from human factors can be mitigated. The case study section reveals that a 50% MPR (market penetration rate) would lead to an 80% risk reduction in specific traffic scenarios, highlighting the potential benefits of the proposed strategy.

### B. CONTRIBUTIONS AND MANUSCRIPT ORGANIZATIONS

The contributions of the manuscript are as follows:

- The dilemma and trilemma scenarios are explicitly defined.
- The polylemma which is a generalization of the dilemma and trilemma, is defined. The polylemma implicate a potential traffic flow risk;
- A human-error-tolerant longitudinal driving strategy is proposed to avoid the polylemma phenomena;
- The ratio and probability of polylemma in current real world traffic flow datasets are calculated;
- The risk reduction and the variation of speed-density after the human-error-tolerant driving is applied are calculated.
- The capacity variation of the proposed strategy is calculated.

The organization of the manuscript is as follows: Section II provides a comprehensive review; Section III explicitly defines the dilemma and trilemma scenarios. In Section IV, the dilemma and trilemma are generalized to the polylemma scenario; Section V calculates the risk reduction and capacity variation of the proposed strategy; Section VI present the case study against several real world dataset; the conclusion and remarks are provided in the last section of the manuscript.

### C. NOTATIONS LIST

RSS: responsibility sensitive safety

HDV and AV: human driven vehicles and autonomous vehicles

AV: autonomous vehicles

IVD: inter-vehicle-distance

MPR: market penetration rate

HET: human error tolerant

PDF and CDF: probability distribution function and cumulative distribution function;

$D_{i\to j}^{RSS}$: the critical distance (or minimal distance) specified by RSS rule between follower (vehicle i) and leader (vehicle j);

$\mathrm{D}^{RSS}$: when all vehicles' speeds and attributes (reaction time, deceleration, acceleration) are the same, then critical distances specified by RSS are also the same, and are indicated by $\mathrm{D}^{RSS}$;

$\mathrm{D}_{i\to i-1}$: the gap between vehicle i and its leader vehicle i-1;

$a_i^{ac}$: the assumed acceleration for vehicle i in the equation calculating $D_{i\to j}^{RSS}$;

$\bar{\mathrm{a}}_i$: the deceleration of vehicle i in the RSS equation;

$\tau_i$: the reaction time of vehicle i in the RSS equation;

$\mathrm{v}_{i,0}$: the initial speed of vehicle i in the RSS equation;

$\bar{\mathrm{a}}_i'$: the moderate deceleration of vehicle i in order to avoid dilemma;

$\bar{\mathrm{a}}_i''$: the moderate deceleration of vehicle i in order to avoid trilemma;

$D_{\mathbf{hw}}^*$ and $\mathbb{D}_{\mathrm{hw}}^*$: the critical distance for dilemma and trilemma avoidance.

$\mathfrak{D}_{\mathrm{hw}}^{\mathrm{X}}$: general expression of the critical distance for general polylemma platoon, where there are X+1 vehicles in the platoon, and the $\mathfrak{D}_{\mathrm{hw}}^{\mathrm{X}}$ is specified for the second vehicle in the platoon. When X=2 (i.e., dilemma), $\mathfrak{D}_{\mathrm{hw}}^{\mathrm{X}}$ is the same as $D_{\mathbf{hw}}^*$; when X=3 (trilemma), $\mathfrak{D}_{\mathrm{hw}}^{\mathrm{X}}$ is the same as $\mathbb{D}_{\mathrm{hw}}^*$;

$\rho$: the market penetration rate of AV.

$\mathrm{l}_{\mathrm{veh}}$: vehicle length.

$\mathrm{f}_{\mathrm{hw}}(\mathrm{h}|\mathrm{v})$ and $\mathrm{F}_{\mathrm{hw}}(\mathrm{h}|\mathrm{v})$: the probability density function and cumulative distribution function of the space gap, which depends on the speed; Note that h doesn't include the vehicle length.

k and v: density and speed.

$\mathrm{p}_{\mathrm{dilemma}}$ and $\mathrm{p}_{\mathrm{trilemma}}$: the occurrence probability of dilemma and trilemma

$\mathrm{p}_{\mathrm{trilemma}\setminus\mathrm{dilemma}}$: probability of a trilemma where there is no dilemma in sub-platoon;

$\mathrm{p}_{reduce|N\ AVs}$: the risk reduction for platoon when there are N consecutive AVs;

$\bar{h}_{\mathrm{dilemma}}$: the average space gap of 3-vehicles platoon which is in critical dilemma state;

$\bar{h}_{trilemma\setminus \text{dilemma}}$: the average space gap of 4-vehicles platoon which is in critical trilemma state while the former three vehicles are not in dilemma;

$p(\bar{h}_{N+1Lemma} = h|v, N\ AVs)$: the probability distribution of average space gap for N+1Lemma when there are N AVs;

$p(\bar{h}_{N+2Lemma\setminus N+1Lemma} = h|v, N\ AVs)$: the probability distribution of average space gap for N+2Lemma when there are N AVs, but the former N+1 vehicles are not in N+1Lemma state;

$p(\bar{h}_{XLemma}|v, \rho)$: the distribution of average space gap for arbitrary polylemma scenario, where X ranges from 2 to infinity;

$p_{HET}(h|v)$ and $p_{HET}(k|v)$: the distribution of the space gap and density when the HET avoidance strategy is applied.

$CDF_{HET}(h|v)$ and $CDF_{HET}(k|v)$: the cumulative distribution function of space gaps and density when the HET is applied.

$CD\,F_{N+2\ Lemma}(h|v, \rho, N\ AVs)$: the CDF of the gaps of N+2Lemma after HET is applied when there are N AVs with MPR as $\rho$;

$k^{eq}_{AA}(v)$: the equilibrium density of autonomous vehicle;

$k^{eq}_{AH}(v)$: the equilibrium density of a AV follows a HDV. It equals $\frac{1}{D^{RSS}(v)+l_{veh}}$

$\mathcal{S}(\cdot \dashv \hbar)$: step function formulation of CDF. It is defined as $\mathcal{S}(h \dashv \hbar) = 0$ if $h < \hbar$, and $\mathcal{S}(h \dashv \hbar) = 1$ otherwise.

Dirac$(\cdot, \cdot)$: shorthand for Dirac function.

$\eta_{\text{dilemma}}$: the ratio of dilemma occurrence in a platoon of real world dataset.

$k_{jam}$: the jam density

$\triangle v_{l-f}(t)$ and $a_f(t)$: the speed difference series and follower's acceleration series for a two vehicle platoon;

$\sigma_{\triangle v_{l-f}, a_f}(\tau)$: cross-covariance function which depends on reaction delay $\tau$;

$\mu_{\triangle v_{l-f}}$ and $\mu_{a_f}$: the mean of the speed difference series and follower's acceleration series

$r_{\triangle v_{l-f}, a_f}(\tau)$: cross-correlation function which depends on reaction delay $\tau$;

$a_{polylemma}$: a predefined acceleration rate for HET;

## II. LITERATURE REVIEW

The adoption of provable safe driving strategies has been a focus in various domains, including microscopic traffic flow modeling [31] and autonomous vehicle development [32]. In the realm of microscopic traffic flow modeling, scholars such as [33], [34], [35], [36] have proposed different formulas that emphasize safety requirements, specifically minimal time headway [37] or space gaps. Moreover, [36] introduces the concept of a perception-reaction process with a pre-determined delay for the follower in the event of an emergency brake initiated by the leading vehicle.

Due to the rise of autonomous vehicles, a set of safety principles has been proposed which are similar to the models developed for microscopic traffic flow analysis [38]. These principles offer more granular details about the underlying movement processes of the involved vehicles [39]. In these safety principles, regulations regarding the inter-vehicle distance [40] or gaps have been introduced. In recent years, the formulation of safety rules for autonomous vehicles has been formalized as a verification problem, following standard practices in the realm of software engineering [41]. An influential formulation in this area is the Responsibility-Sensitive Safety framework, developed by Intel and Mobileye [21]. Considering longitudinal driving, RSS specifies a minimum distance between two vehicles. This is accomplished by assuming a maximum deceleration for the leading vehicle and a normal deceleration for the following vehicle.

The original formulation of the RSS for longitudinal driving is relatively simple and relies on assumptions that may be too strong. Many critics have proposed improvements to address these issues. For example, one criticism is that the calculated distance is overly conservative [42], compromising efficiency. Remedies have been developed to address this, such as the establishment of a situation awareness RSS [43], which distinguishes different traffic states. Another improvement is the development of a differentiated RSS triggering algorithm [44], which calculates three RSS distances based on different speed profiles and deceleration rates of the following vehicle. Similar improvements in conservation can be found in other studies [42]. Additionally, when autonomous vehicles are equipped with communication capabilities, such as in ACC [37], [45] and CACC configurations [40], or additional sensors [46], the RSS can be more specific and the results can be further investigated. Furthermore, the parameters in the RSS models are calibrated using safety critical data extracted from normal driving observations [47], [48].

The aforementioned theoretical and empirical studies indicate that incorporating safety by design would be advantageous for the advancement of autonomous vehicles [49]. However, current approaches primarily concentrate on the mathematical modeling of AVs' responsibilities, assuming that human-driven vehicles will behave as expected [50]. In a mixed flow of autonomous and human-driven vehicles, HDVs may exhibit considerable heterogeneity [51] and significant risks, which cannot be disregarded by AVs as they could potentially lead to unavoidable collisions [52].

Given that HDVs are not controllable directly, it is feasible to achieve specific objectives through AVs' controllers. Utilizing controlled vehicles to influence the movement dynamics is not a novel concept [53]. For example, AVs have been employed to mitigate jerks or oscillation waves [28], [54]. However, the utilization of AVs to mitigate safety risks in RSS configurations has not yet been reported. This research examines scenarios where HDVs themselves do not adhere to the rules of safe driving and pose risks to AVs. A novel concept called polylemma, which generalizes the notions of dilemma and trilemma, is proposed to describe situations where AVs overlook risks due to HDVs' violations of safety rules. Polylemma is triggered when one or more HDVs violate the RSS, placing the nearest AVs in a precarious condition. An associated avoidance strategy is put forth to address this issue.

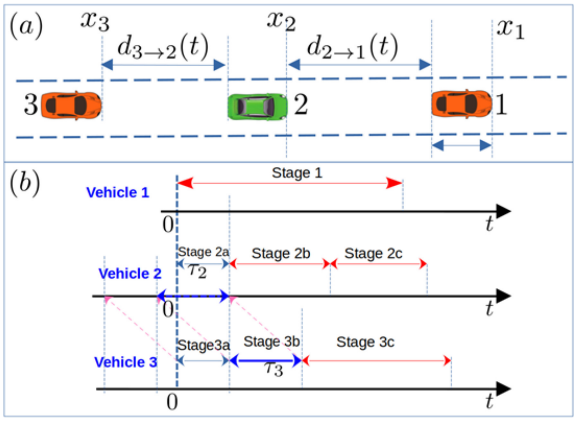


**FIGURE 1.** Time line of the deceleration events for dilemma description.

## III. DILEMMA AND TRILEMMA: DEFINITION AND AVERAGE HEADWAY

### A. DEFINITION OF DILEMMA

Consider a platoon consisting of three vehicles denoted as 1, 2, and 3.

Fig. 1-a provides the initial configuration, where the positions of the vehicles are denoted as $x_1$, $x_2$ and $x_3$ respectively. Initial speeds are $v_{1,0}$, $v_{2,0}$ and $v_{3,0}$ respectively. At time t=0, vehicle 1 undergoes an abrupt deceleration, reaching a full stop with a maximum deceleration represented by $\bar{a}_1$. Following a duration of $\tau_2$ (corresponding to the reaction delay of vehicle 2), vehicle 2 initiates deceleration with a value of $\bar{a}_2$. Subsequently, after a time interval of $\tau_3$ (characterizing the reaction delay of vehicle 3), vehicle 3 decelerates to a complete stop using a deceleration rate of $\bar{a}_3$.

The RSS (Responsibility-Sensitive Safety) framework defines a minimum inter-vehicle distance for vehicle 2, denoted as $D^{RSS}_{2\to1}$, between vehicle 1 and 2. This distance guarantees that, throughout the deceleration process, there will be no collisions between the two vehicles, regardless of the control policy implemented by vehicle 2. It is assumed that during the reaction time of $\tau_2$, vehicle 2 may utilize the acceleration $a_2^{ac}$. The calculation of this minimum distance (i.e., $D^{RSS}_{2\to1}$) is outlined in [45] as follows. Suppose during the reaction time of $\tau_2$, vehicle 2's acceleration is $a_2^{ac}$, then if $\bar{a}_1 > \bar{a}_2$,

$$D^{RSS}_{2\to1} = \left[\frac{a_2^{ac}}{2}\left(\frac{a_2^{ac}}{\bar{a}_2}+1\right)\tau_2^2 + v_{2,0}\left(\frac{a_2^{ac}}{\bar{a}_2}+1\right)\tau + \left(\frac{(v_{2,0})^2}{2\bar{a}_2} - \frac{(v_{1,0})^2}{2\bar{a}_1}\right)\right]_+ \tag{1}$$

If $\bar{a}_1 < \bar{a}_2$:

$$D^{RSS}_{2\to1} = \begin{cases} (v_{2,0}-v_{1,0})\tau_2 + \frac{\tau_2^2}{2}(\bar{a}_1+a_2^{ac}) + \frac{(v_{1,0}-v_{2,0}-(\bar{a}_1+a_2^{ac})\tau_2)^2}{2(\bar{a}_2-\bar{a}_1)}, \\ \text{if } \tau_2 \in \left[\frac{v_{1,0}-v_{2,0}}{a_2^{ac}+\bar{a}_1}, \frac{v_{1,0}-\frac{\bar{a}_2}{\bar{a}_1}v_{2,0}}{a_2^{ac}+\bar{a}_2}\right] \\ \left[\frac{a_2^{ac}}{2}\left(\frac{a_2^{ac}}{\bar{a}_2}+1\right)\tau_2^2 + v_{2,0}\left(\frac{a_2^{ac}}{\bar{a}_2}+1\right)\tau_2 + \left(\frac{(v_{2,0})^2}{2\bar{a}_2} - \frac{(v_{1,0})^2}{2\bar{a}_1}\right)\right]_+, \text{otherwise} \end{cases} \tag{2}$$

Similarly, we can write the RSS distance of vehicle 3, $D^{RSS}_{3\to2}$. If vehicle 3 violates the RSS rule by maintaining a gap $d_{3\to2}$ smaller than $D^{RSS}_{3\to2}$ from vehicle 2, a collision may still occur between vehicle 3 and 2. To avoid a collision, vehicle 2 can adopt a moderate deceleration $\bar{a}'_2$ instead of $\bar{a}_2$. However, a moderate deceleration of vehicle 2 implies that the distance $D^{RSS}_{2\to1}$ is insufficient, and a greater distance which accommodate $\bar{a}'_2$, indicated as $D^*_{\mathbf{hw}}$, is required. Hence we have the definition of dilemma as follows:

**Definition of dilemma**. If the actual gap $d_{2\to1}$ satisfies the condition $D^{RSS}_{2\to1} < d_{2\to1} < D^*_{\mathbf{hw}}$, the following conclusions can be drawn: 1) If vehicle 1 decelerates abruptly, a collision between vehicle 2 and vehicle 1 can be avoided if vehicle 2 use the deceleration of $\bar{a}_2$. However, vehicle 2 will still collide with vehicle 3, who violates the RSS rule. 2) To avoid a collision with vehicle 3, vehicle 2 must apply deceleration no stronger than $\bar{a}'_2$. However, the condition $d_{2\to1} < D^*_{hw}$ implies that a collision will occur between vehicle 2 and vehicle 1. This explanation clarifies the situation of vehicle 2, where at least one collision is inevitable, thus defining it as a dilemma. The events classification for $d_{1\to2}$ is provided in (3).

$$\begin{cases} d_{1\to2} > D^*_{\mathbf{hw}}, & no\ RSS\ and\ no\ Dilemma \\ D^{RSS}_{2\to1} \le d_{1\to2} < D^*_{\mathbf{hw}}, & Dilemma \\ d_{1\to2} < D^{RSS}_{2\to1}, & RSS\ violation \end{cases} \tag{3}$$

The formula for the critical dilemma gap $D^*_{\mathbf{hw}}$ can be derived as follows, which is divided into two steps: 1) In the first step, we calculate the required deceleration $\bar{a}'_2$ when the real-time gap $d_{3\to2} < D^{RSS}_{3\to2}$, indicating a violation of RSS by vehicle 3; 2) Given the deceleration $\bar{a}'_2$, we determine the value of $D^*_{\mathbf{hw}}$.

Step 1: solving required moderate deceleration $\bar{\mathbf{a}}'_2$

Essentially, determining the deceleration $\bar{a}'_2$ involves solving the inverse problem of RSS. The original expression [45] for the RSS distance $D^{RSS}_{3\to2}$ is expressed as follows.

If

$$\{\bar{a}_2 \ge \bar{a}_3\} \cup \left\{\{\bar{a}_2 < \bar{a}_3\} \cap \left\{\tau \notin \left[\frac{v_{2,0}-v_{3,0}}{a_3^{ac}+\bar{a}_2}, \frac{v_{2,0}\frac{\bar{a}_3}{\bar{a}_2}-v_{3,0}}{a_3^{ac}+\bar{a}_2}\right]\right\}\right\}$$

$$D^{RSS}_{3\to2} = \left[\frac{a_3^{ac}}{2}\left(\frac{a_3^{ac}}{\bar{a}_3}+1\right)\tau_3^2 + v_{3,0}\left(\frac{a_3^{ac}}{\bar{a}_3}+1\right)\tau_3 + \left(\frac{(v_{3,0})^2}{2\bar{a}_3} - \frac{(v_{2,0})^2}{2\bar{a}_2}\right)\right]_+ \tag{4}$$

Else:

$$D^{RSS}_{3\to2} = (v_{3,0}-v_{2,0})\tau_3 + \frac{\tau_3^2}{2}(\bar{a}_2+a_3^{ac}) + \frac{(v_{2,0}-v_{3,0}-(\bar{a}_2+a_3^{ac})\tau_3)^2}{2(\bar{a}_3-\bar{a}_2)} \tag{5}$$

We replace $D^{RSS}_{3\to2}$ in (4) and (5) with the given gap $D_{3\to2}$, and solve the deceleration $\bar{a}_2$. The candidate solutions of $\bar{a}'_2$

from (4) and (5) (which are indicated using $\bar{a}_2^{\#}$ and $\bar{a}_2^{\odot}$) are as follows:

$$\begin{cases} \bar{a}_2^{\#} = \dfrac{\bar{a}_3\, v_{2,0}^2}{(a_3^{ac})^2\tau_3^2 + a_3^{ac}\bar{a}_3\tau_3^2 + 2a_3^{ac}\tau_3\, v_{3,0} - 2d_{3\to 2}\bar{a}_3 + 2\bar{a}_3\tau_3\, v_{3,0} + v_{3,0}^2} \quad (a) \\ \qquad -(a_3^{ac})^2\tau_3^2 - a_3^{ac}\bar{a}_3\tau_3^2 + 2a_3^{ac}\tau_3\, v_{2,0} - 2a_3^{ac}\tau_3\, v_{3,0} - \\ \bar{a}_2^{\odot} = \dfrac{2d_{3\to 2}\bar{a}_3 - 2\bar{a}_3\tau_3\, v_{2,0} + 2\bar{a}_3\tau_3\, v_{3,0} - v_{2,0}^2 + 2\, v_{2,0}\, v_{3,0} - v_{3,0}^2}{a_2^{ac}\tau_3^2 - 2d_{3\to 2} + \bar{a}_3\tau_3^2 - 4\tau_3\, v_{2,0} + 4\tau_3\, v_{3,0}} \quad (b) \end{cases} \tag{6}$$

Then, if

$$\{\bar{a}_2^{\#} \geq \bar{a}_3\} \cup \left\{ \{\bar{a}_2^{\#} < \bar{a}_3\} \cap \left\{ \tau_3 \notin \left[ \frac{v_{2,0} - v_{3,0}}{a_3^{ac} + \bar{a}_2^{\#}}, \frac{v_{2,0}\frac{\bar{a}_3}{\bar{a}_2^{\#}} - v_{3,0}}{a_3^{ac} + \bar{a}_2^{\#}} \right] \right\} \right\}$$

$$\begin{aligned} \bar{a}_2' &= \bar{a}_2^{\#} \\ &= \frac{\bar{a}_3\, v_{2,0}^2}{(a_3^{ac})^2\tau_3^2 + a_3^{ac}\bar{a}_3\tau_3^2 + 2a_3^{ac}\tau_3\, v_{3,0} - 2d_{3\to 2}\bar{a}_3 + 2\bar{a}_3\tau_3\, v_{3,0} + v_{3,0}^2} \end{aligned} \tag{7}$$

Else:

$$\begin{aligned} \bar{a}_2' &= \bar{a}_2^{\odot} \\ &= \frac{\begin{array}{c} -(a_3^{ac})^2\tau_3^2 - a_3^{ac}\bar{a}_3\tau_3^2 + 2a_3^{ac}\tau_3\, v_{2,0} - 2a_3^{ac}\tau_3\, v_{3,0} - \\ 2d_{3\to 2}\bar{a}_3 - 2\bar{a}_3\tau_3\, v_{2,0} + 2\bar{a}_3\tau_3\, v_{3,0} - v_{2,0}^2 + 2\, v_{2,0}\, v_{3,0} - v_{3,0}^2 \end{array}}{a_2^{ac}\tau_3^2 - 2d_{3\to 2} + \bar{a}_3\tau_3^2 - 4\tau\, v_{2,0} + 4\tau_3\, v_{3,0}} \end{aligned} \tag{8}$$

Step 2: find the $D_{\mathrm{hw}}*$ given $\bar{\mathbf{a}}_2'$

The problem of finding $D_{\mathrm{hw}}^*$ is the same as in conventional RSS setting. The only difference is that the original deceleration of vehicle 2 (i.e., $\bar{a}_2$) is replaced by the required deceleration, $\bar{a}_2'$. The result is summarized in the following.

If

$$\{\bar{a}_1 \geq \bar{a}_2'\} \cup \left\{ \{\bar{a}_1 < \bar{a}_2'\} \cap \left\{ \tau_2 \notin \left[ \frac{v_{1,0} - v_{2,0}}{a_2^{ac} + \bar{a}_1}, \frac{v_{2,0}\frac{\bar{a}_2'}{\bar{a}_1} - v_{2,0}}{a_2^{ac} + \bar{a}_1} \right] \right\} \right\}:$$

$$D_{\mathbf{hw}}^* = \left[ \frac{a_2^{ac}}{2}\left(\frac{a_2^{ac}}{\bar{a}_2'} + 1\right)\tau_2^2 + v_{2,0}\left(\frac{a_2^{ac}}{\bar{a}_2'} + 1\right)\tau + \left(\frac{(v_{2,0})^2}{2\bar{a}_2'} - \frac{(v_{1,0})^2}{2\bar{a}_1}\right) \right]_+ \tag{9}$$

Else:

$$\begin{aligned} D_{\mathbf{hw}}^* = {} & (v_{2,0} - v_{1,0})\tau_2 + \frac{\tau_2^2}{2}(\bar{a}_1 + a_2^{ac}) \\ & + \frac{(v_{1,0} - v_{2,0} - (\bar{a}_1 + a_2^{ac})\tau_2)^2}{2(\bar{a}_2' - \bar{a}_1)} \end{aligned} \tag{10}$$

The results of (7) and (8) can be used to expand (9) and (10), but due to their complexity, they will not be presented here. The difference between the RSS distance $D_{2\to 1}^{RSS}$ and $D_{\mathbf{hw}}^*$ represents the extra distance reserved for the reckless driving of vehicle 3.

#### 1) STATE SPACE OF RSS

Using $D_{\mathbf{hw}}^*$, the state space comprising two gaps ($d_{3\to 2}$ and $D_{2\to 1}$) can be partitioned into five zones, as presented in Fig. 2. In zone 1 (upper left zone), vehicle 3 adheres to the recommended safe separation distance, while vehicle 2

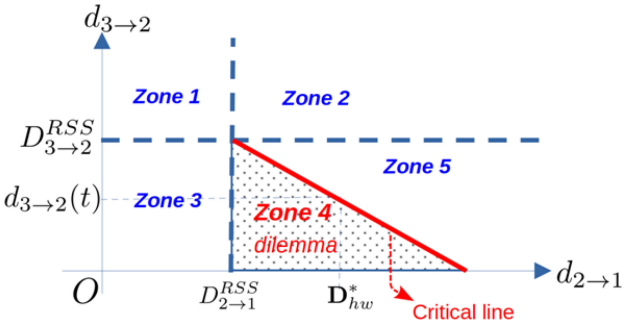


**FIGURE 2.** Dilemma specification.

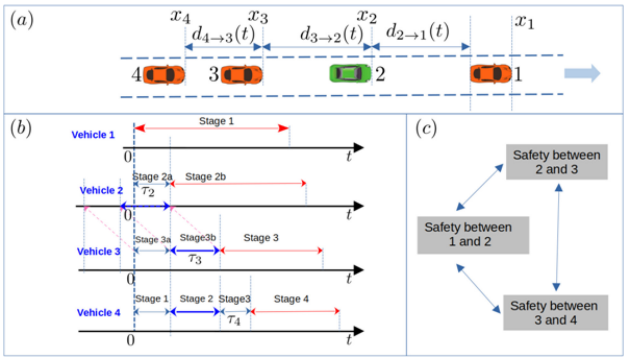


**FIGURE 3.** Trilemma scenario: (a) the example platoon; (b) time line of the deceleration; (c) the trilemma.

violates it. In zone 2, both vehicles (2 and 3) abide by the RSS guidelines. In zone 3, both vehicles contravene the RSS. In zone 4, vehicle 3 maintains a shorter gap (less than $D_{3\to 2}^{RSS}$) from vehicle 2, necessitating a gentle deceleration from vehicle 2 to ensure an additional buffer distance, represented by the critical line. Each gap value $D_{3\to 2}(t)$ for vehicle 3 has a corresponding critical gap for vehicle 2 ($D_{\mathbf{hw}}^*$ in the figure). When the actual gap is smaller than $D_{\mathbf{hw}}^*$, vehicle 2 is at risk of colliding either with vehicle 1 or vehicle 3.

### B. TRILEMMA: A PLATOON WITH THREE VEHICLES

The trilemma is an extension of the dilemma and can be illustrated in Fig. 3-a with four vehicles (numbered 1 to 4). In this scenario, assuming vehicles 2 and 3 follow the recommended safe separation, and vehicle 4 violates the RSS. Initial speeds are $v_{1,0}$, $v_{2,0}$, $v_{3,0}$ and $v_{4,0}$ respectively. If vehicle 1 abruptly decelerates, a collision between vehicle 4 and vehicle 3 would occur.

To prevent a collision with vehicle 4, vehicle 3 is assumed to use a gentle deceleration. Consequently, vehicle 2 (an AV) also needs to decelerate moderately since it cannot directly control vehicle 3. This necessitates maintaining a special distance between vehicles 2 and 1 to accommodate vehicle 3's deceleration behavior. This required distance is denoted as $\mathbb{D}_{\mathrm{hw}}^*$, and is dependent on various factors within the platoon, such as $\tau_2, \tau_3, \tau_4, v_{1,0}, v_{2,0}, v_{3,0}, v_{4,0}, d_{4\to 3}, d_{3\to 2}$. This situation is represented by the concept of the trilemma,

visualized in Fig. 3-c. The trilemma showcases three triangles: one representing the safety between vehicle 1 and 2, another representing the safety between 2 and 3, and the third representing the safety between 4 and 3. Trilemma indicates that, one cannot guarantee the above three safeties simultaneously, if $d_{2\to1} < \mathbb{D}^*_{\text{hw}}$. Thus there must at least one crash event occurs.

The determination of $\mathbb{D}^*_{\text{hw}}$ involves several steps, akin to the dilemma scenario. Firstly, we compute the required deceleration $\bar{a}'_3$ for vehicle 3 to adjust for vehicle 4's error, considering the parameters $d_{4\to3}$ and $\tau_4$. Secondly, based on $\bar{a}'_3$ and $\tau_3$, we determine the deceleration $\bar{a}''_2$ for vehicle 2, which accounts for accommodating vehicle 3. Lastly, we calculate the $\mathbb{D}^*_{\text{hw}}$.

Step 1: find the required deceleration $\bar{a}'_2$

The process of the deriving of $\bar{a}'_3$ is the same as $\bar{a}'_2$ in (7) and (8). The process is omitted and the results are presented directly. The candidate solution ($\bar{a}^{\#}_3$ and $\bar{a}^{\odot}_3$, basically the same as $\bar{a}^{\#}_2$ and $\bar{a}^{\odot}_2$ in (6) of desired deceleration of vehicle 3 from RSS equation are:

$$\begin{cases} \bar{a}^{\#}_3 = \dfrac{\bar{a}_4\, v^2_{3,0}}{(a^{ac}_4)^2\tau_4^2+a^{ac}_4\bar{a}_4\tau_4^2+2a^{ac}_4\tau_4\, v_{4,0}-2d_{4\to3}\bar{a}_4+2\bar{a}_4\tau_4\, v_{4,0}+v^2_{4,0}} \quad (a) \\ \qquad -(a^{ac}_4)^2\tau_4^2 - a^{ac}_4\bar{a}_4\tau_4^2 + 2a^{ac}_4\tau_4\, v_{3,0} - 2a^{ac}_4\tau_4\, v_{4,0} - \\ \bar{a}^{\odot}_3 = \dfrac{2d_{4\to3}\bar{a}_4 - 2\bar{a}_4\tau_4\, v_{3,0} + 2\bar{a}_4\tau_4\, v_{4,0} - v^2_{3,0} + 2\, v_{3,0}\, v_{4,0} - v^2_{4,0}}{a^{ac}_3\tau_4^2 - 2d_{4\to3} + \bar{a}_4\tau_4^2 - 4\tau_4\, v_{3,0} + 4\tau_4\, v_{4,0}} \quad (b) \end{cases} \tag{11}$$

Then, if

$$\{\bar{a}^{\#}_3 \geq \bar{a}_4\} \cup \left\{ \{\bar{a}^{\#}_3 < \bar{a}_4\} \cap \left\{ \tau_4 \notin \left[ \frac{v_{3,0}-v_{4,0}}{a^{ac}_4+\bar{a}^{\#}_3}, \frac{v_{3,0}\frac{\bar{a}_4}{\bar{a}^{\#}_3}-v_{4,0}}{a^{ac}_4+\bar{a}^{\#}_3} \right] \right\} \right\}$$

$$\bar{a}'_3 = \frac{\bar{a}_4\, v^2_{3,0}}{(a^{ac}_4)^2\tau_4^2 + a^{ac}_4\bar{a}_4\tau_4^2 + 2a^{ac}_4\tau_4\, v_{4,0} - 2d_{4\to3}\bar{a}_4 + 2\bar{a}_4\tau_4\, v_{4,0} + v^2_{4,0}} \tag{12}$$

Else, the required deceleration is calculated as:

$$\bar{a}'_3 = \frac{-(a^{ac}_4)^2\tau_4^2 - a^{ac}_4\bar{a}_4\tau_4^2 + 2a^{ac}_4\tau_4\, v_{3,0} - 2a^{ac}_4\tau_4\, v_{4,0} - 2d_{4\to3}\bar{a}_4 - 2\bar{a}_4\tau_4\, v_{3,0} + 2\bar{a}_4\tau_4\, v_{4,0} - v^2_{3,0} + 2\, v_{3,0}\, v_{4,0} - v^2_{4,0}}{a^{ac}_3\tau_4^2 - 2d_{4\to3} + \bar{a}_4\tau_4^2 - 4\tau_4\, v_{3,0} + 4\tau_4\, v_{4,0}} \tag{13}$$

Step 2: find the $\bar{a}''_2$

The solution method of $\bar{a}''_2$ basically is the same as $\bar{a}'_3$. We directly present the results. When derive $\bar{a}''_2$, it is assumed that the vehicle 3 would decelerate using $\bar{a}'_3$, rather than original $\bar{a}_3$. Thus we just need to replace $\bar{a}_3$ using $\bar{a}'_3$ in (7) and (8), and get the results:

$$\bar{a}''_2 = \frac{\bar{a}'_3 \cdot v^2_{2,0}}{(a^{ac}_3)^2\tau_3^2 + a^{ac}_3 \cdot \bar{a}'_3 \cdot \tau_3^2 + 2a^{ac}_3\tau_3\, v_{3,0} - 2d_{3\to2} \cdot \bar{a}'_3 + 2\bar{a}'_3 \cdot \tau_3\, v_{3,0} + v^2_{3,0}} \tag{14}$$

Or

$$\bar{a}''_2 = \frac{-(a^{ac}_3)^2\tau_3^2 - a^{ac}_3 \cdot \bar{a}'_3 \cdot \tau_3^2 + 2a^{ac}_3\tau_3\, v_{2,0} - 2a^{ac}_3\tau_3\, v_{3,0} - 2d_{3\to2} \cdot \bar{a}'_3 - 2\bar{a}'_3 \cdot \tau_3\, v_{2,0} + 2\bar{a}'_3 \cdot \tau_3\, v_{3,0} - v^2_{2,0} + 2\, v_{2,0}\, v_{3,0} - v^2_{3,0}}{a^{ac}_2\tau_3^2 - 2d_{3\to2} + \bar{a}'_3 \cdot \tau_3^2 - 4\tau\, v_{2,0} + 4\tau_3\, v_{3,0}} \tag{15}$$

Similarly we also didn't to expand (14) and (15) using (12) and (13) as the results are complex.

Step 3: find the $\mathbb{D}_{\text{hw}}*$

The solution of $D^*_{\mathbf{hw}}$ basically is similar to $D^*_{\mathbf{hw}}$. We replace $\bar{a}'_2$ in (9) and (10) using $\bar{a}''_2$, and we have the results. If

$$\{\bar{a}_1 \geq \bar{a}''_2\} \cup \left\{ \{\bar{a}_1 < \bar{a}''_2\} \cap \left\{ \tau_2 \notin \left[ \frac{v_{1,0}-v_{2,0}}{a^{ac}_2+\bar{a}_1}, \frac{v_{2,0}\frac{\bar{a}''_2}{\bar{a}_1}-v_{3,0}}{a^{ac}_2+\bar{a}_1} \right] \right\} \right\}:$$

$$D^*_{\mathbf{hw}} = \left[ \frac{a^{ac}_2}{2}\left(\frac{a^{ac}_2}{\bar{a}''_2}+1\right)\tau_2^2 + v_{2,0}\left(\frac{a^{ac}_2}{\bar{a}''_2}+1\right)\tau + \left(\frac{(v_{2,0})^2}{2\bar{a}''_2} - \frac{(v_{1,0})^2}{2\bar{a}_1}\right) \right]_+ \tag{16}$$

Else, the critical dilemma distance is specified by:

$$D^*_{\mathbf{hw}} = (v_{2,0}-v_{1,0})\tau_2 + \frac{\tau_2^2}{2}(\bar{a}_1+a^{ac}_2) + \frac{(v_{1,0}-v_{2,0}-(\bar{a}_1+a^{ac}_2)\tau_2)^2}{2(\bar{a}''_2-\bar{a}_1)} \tag{17}$$

## IV. POLYLEMMA: GENERALIZATION OF DILEMMA AND A HUMAN ERROR TOLERANT SOLUTION

### A. POLYLEMMA DEFINITION

Consider a platoon consisting of N vehicles, where N can vary from 3 to infinity. The leader vehicle designated as vehicle 1. If the Nth vehicle in the platoon violates the rules of the Responsibility-Sensitive Safety and comes too close to vehicle N-1, it necessitates a more moderate deceleration by vehicle N-1 to avoid a collision. This, in turn, requires a gentle deceleration by vehicle N-2, and so on. The aforementioned deductive chain can be iterated until it reaches vehicle 2, where a larger gap between vehicle 2 and vehicle 1 is necessary to accommodate the moderate decelerations of vehicles 3 to N-1. Once vehicle 2 fails to maintain such gap, polylemma forms. Theoretically, vehicle number X ranges from 2 to infinity. A platoon adhering to the polylemma consists of X+1 vehicles, with the critical gap between vehicle 1 (the platoon leader) and vehicle 2 denoted as $\mathfrak{D}^{\text{X}}_{\text{hw}}$.

### B. RECURSIVE INFLUENCE OF FOLLOWER'S DECELERATION CHANGE ON LEADER'S DECELERATION UNDER THE CRITICAL RSS SCENARIO

Downstream vehicles must undergo a moderate deceleration to accommodate the decreasing mild decelerations of upstream vehicles. We examine the variation in deceleration of downstream vehicles in response to decelerations of upstream vehicles within the critical RSS gap. Considering two vehicles, with vehicle 1 as the leader and vehicle 2 as the follower, we derive the differential $\frac{\partial \bar{a}_1}{\partial \bar{a}_2}$ under RSS requirement. If $\{\bar{a}_1 \geq \bar{a}_2\} \cup \{\{\bar{a}_1 < \bar{a}_2\} \cap \{\tau \notin [\frac{v_{1,0}-v_{2,0}}{a^{ac}_2+\bar{a}_1}, \frac{v_{1,0}\frac{\bar{a}_2}{\bar{a}_1}-v_{2,0}}{a^{ac}_2+\bar{a}_1}]\}\}$:

$$\frac{\partial \bar{a}_1}{\partial \bar{a}_2} = \frac{0.25\bar{a}_2 \cdot v^2_{1,0}\left(-a^{ac}_2\tau_2^2 + 2D^{RSS}_{2\to1} - 2\tau_2 D^{RSS}_{2\to1}\right)}{\left(0.5(a^{ac}_2)^2\tau_2^2 + 0.5a^{ac}_2\bar{a}_2\tau_2^2 + a^{ac}_2\tau_2\, v_{2,0} - D^{\text{RSS}}_{2\to1}\bar{a}_2 + \bar{a}_2\tau_2\, v_{2,0} + v^2_{2,0}\right)^2}$$

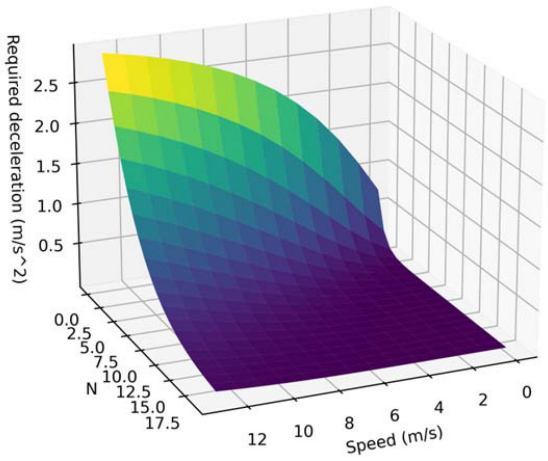


**FIGURE 4.** Recursive influence of required deceleration in a platoon of 20 vehicles.

$$+\frac{v_{1,0}^2}{(a_2^{ac})^2\tau_2^2+a_2^{ac}\bar{a}_2\tau_2^2+2a_2^{ac}\tau_2\, v_{2,0}-2\,D_{2\to1}^{RSS}\,\bar{a}_2+2\bar{a}_2\tau_2\, v_{2,0}+v_{2,0}^2} \quad (18)$$

Else:

$$\frac{\partial\bar{a}_1}{\partial\bar{a}_2}=-\frac{\tau_2^2\left(\begin{array}{c}-(a_2^{ac})^2\tau_2^2-a_2^{ac}\bar{a}_2\tau_2^2+2a_2^{ac}\tau_2\, v_{1,0}-2a_2^{ac}\tau_2\, v_{2,0}-\\ 2D_{2\to1}^{RSS}\bar{a}_2-2\bar{a}_2\tau_2\, v_{1,0}+2\bar{a}_2\tau_2\, v_{2,0}-v_{1,0}^2+2\, v_{10}\, v_{2,0}-v_{2,0}^2\end{array}\right)}{\left(a_2^{ac}\tau_2^2-2D_{2\to1}^{RSS}+\bar{a}_2\tau_2^2-4\tau_2\, v_{1,0}+4\tau_2\, v_{2,0}\right)^2}$$
$$+\frac{-a_2^{ac}\tau_2^2-2D_{2\to1}^{RSS}-2\tau_2\, v_{1,0}+2\tau_2\, v_{2,0}}{a_2^{ac}\tau_2^2-2D_{2\to1}^{RSS}+\bar{a}_2\tau_2^2-4\tau_2\, v_{1,0}+4\tau_2\, v_{2,0}} \quad (19)$$

Fig. 4 presents the required decelerations for each vehicle in the 20-vehicle platoon. Interestingly, the second-to-last vehicle needs to apply a deceleration of $-2.9\text{m/s}^2$, as opposed to the originally prescribed $-3.5\text{sm/s}^2$, at the speed of 12 m/s. Consequently, the third-to-last vehicle must apply a deceleration of $-2.497\text{m/s}^2$. This process continues, eventually leading to the first vehicle in the platoon requiring a deceleration of $-0.06\text{m/s}^2$, significantly lower than the initial $-3.5\text{m/s}^2$.

### C. HUMAN ERROR TOLERANT (HET) STRATEGY

The HET strategy involves three steps: identification of the violating vehicle in terms of RSS, calculation of the necessary deceleration and gap for each downstream vehicle until the RSS-specified gap is satisfied, and determination of which autonomous vehicle in the polylemma should implement the HET strategy. Finally, the execution of the HET strategy takes place.

#### 1) POLYLEMMA-PLATOON IDENTIFICATION

In a platoon, simultaneous violations of the RSS may occur in multiple vehicles. Each violation creates micro-Lemma instances. When a micro-Lemma requires the adjustment of required deceleration in another micro-Lemma downstream, the two micro-Lemma instances combine to form a macro-Lemma. Fig. 5 illustrates a scenario involving a micro-Lemma platoon and a Macro-Lemma platoon:

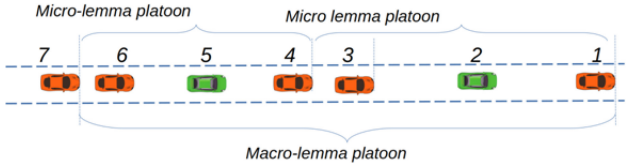


**FIGURE 5.** Lemma-platoon.

vehicle 7 violates the RSS rule, necessitating behavioral adjustments from vehicle 5 (an AV) and vehicle 4 (a HDV) to accommodate the violation. Consequently, vehicles 4 to 7 form a trilemma. Additionally, vehicle 4 itself violates the RSS rule, resulting in another trilemma formed by vehicles 1 to 4.

#### 2) HUMAN ERROR TOLERANT (HET) BY GAP RESERVATION

Regarding the scenarios in Fig. 5, there are two options:

Option 1: Vehicle 2 addresses the avoidance of the Trilemma by solely taking into account the RSS violation of vehicle 4, while vehicle 5 focuses on the avoidance of the Dilemma by considering only the RSS violation of vehicle 7. The critical gaps ($\mathfrak{D}_{\text{hw}}^3$, or equivalently $\mathbb{D}_{\text{hw}}^*$) for these two vehicles are calculated separately.

Option 2: Vehicle 5 remains inactive, while vehicle 2 implements the avoidance of the Hexalemma (which involves 7 vehicles) by considering the RSS violations of both vehicle 4 and vehicle 7. In this case, the critical gap is denoted as $\mathfrak{D}_{\text{hw}}^6$, as indicated in general for polylemma avoidance (where X equals 6).

Based on the findings from Fig. 4, which show that a milder deceleration is required when more vehicles are involved, we select option 1 as the preferred avoidance strategy for polylemma. Vehicle 2 and vehicle 5 need to increase the gaps to a value denoted as $\mathbb{D}_{\text{hw}}^*$.

## V. POTENTIAL RISK REDUCTION, FUNDAMENTAL DIAGRAM AND CAPACITY IMPLICATIONS

The implementation of the polylemma-avoidance strategy has the potential to mitigate risks and alter the time headway. In this section, we focus on quantifying the impact of this strategy on risk reduction, changes in the average gap, and capacity.

### A. GENERAL STEPS

Our analysis primarily focuses on longitudinal driving in the context of background traffic flow. To calculate the risk reduction and capacity after the HET strategy is applied, we use the following steps: 1) Firstly, the background traffic flow is represented by density distribution for a given speed v, $p(k|v)$; 2) the probability of the polylemma scenario for each possible X is formulated, considering a given AV's MPR $\rho$ and speed v; 3) the risk reduction probability and density k for any speed v following the application of the HET strategy are formulated. The resultant

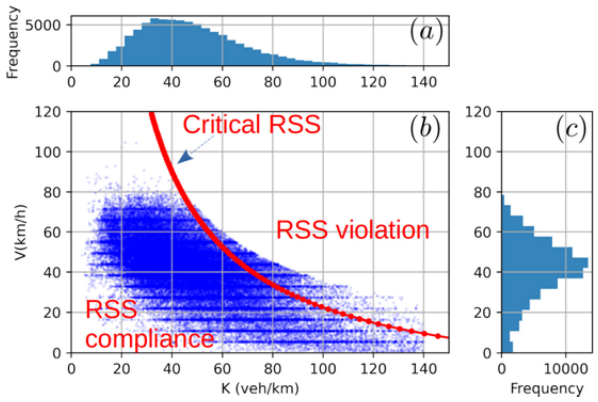


**FIGURE 6.** Background traffic flow representation: NGSIM dataset.

density and gap distributions are denoted as $\mathrm{p_{HET}}(k|v)$ and $\mathrm{p_{HET}}(h|v)$, respectively. Their cumulative distribution functions are represented as $\mathrm{CDF_{HET}}(k|v)$ and $\mathrm{CDF_{HET}}(h|v)$. 4) the capacity is determined as the maximum expected flow rate.

Regarding background traffic flow in step 1, taking the U.S. background traffic flow as an example, the speed-density relationship in the NGSIM dataset is illustrated in Fig. 12-b (using Edie definition). The RSS rule defined in (1) and (2) establishes a relationship between speed and space headway which is indicated by critical RSS in Fig. 6-b. The probability distribution of speeds is represented by p(v) (shown in Fig. 6-c), while the conditional distribution of density given a speed is denoted as p(k|v). With the assumption of a known and uniform vehicle length ($\mathrm{l_{veh}}$), we can indirectly calculate the distribution of gaps based on the relationship between density (k) and gap (h), i.e., $\mathrm{k} = \frac{1}{\mathrm{h}+\mathrm{l_{veh}}}$. We use the notations $\mathrm{f_{hw}}(\mathrm{h})$ (or $\mathrm{f_{hw}}(\mathrm{h}|\mathrm{v})$) and $\mathrm{F_{hw}}(\mathrm{h})$ (or $\mathrm{F_{hw}}(\mathrm{h}|\mathrm{v})$) to represent the probability density function PDF (or conditional PDF) and cumulative distribution function CDF (or conditional CDF) of space gaps in the background traffic flow.

### B. POTENTIAL RISK REDUCTION

We begin by examining the scenarios involving a single AV, namely the dilemma (represented in Fig. 1) and trilemma (shown in Fig. 3). In the dilemma, the vehicle types sequence is H→A→H, while in the trilemma, it is H→H→A→H, with the arrow indicating the direction from follower to leader. The probabilities of these two combinations, given a specific MPR $\rho$, are calculated as follows:

$$\begin{cases} p(HAH) = (1-\rho)\rho(1-\rho) = \rho(1-\rho)^2 & (a) \\ p(HHAH) = (1-\rho)(1-\rho)\rho(1-\rho) = \rho(1-\rho)^3 & (b) \end{cases} \tag{20}$$

The probability of the dilemma can be calculated by multiplying the probability of the combination occurring with the probability that the gap $d_{3\to2}$ between the 2nd and 3rd vehicles is smaller than the critical RSS:

$$\begin{aligned} \mathrm{p_{dilemma}} &= p(HAH)\cdot p\left(d_{3\to2} < D_{3\to2}^{RSS}\right) \\ &= \rho(1-\rho)^2\cdot p\left(d_{3\to2} < D_{3\to2}^{RSS}\right) \end{aligned} \tag{21}$$

The probability $p(d_{3\to2} < D_{3\to2}^{RSS})$ implies that vehicle 3 already violate the RSS. The probability of trilemma is calculated similarly:

$$\begin{aligned} \mathrm{p_{trilemma}} &= p(HHAH)\cdot p\left(d_{4\to3} < D_{4\to3}^{RSS}\right)\cdot p\left(d_{3\to2} < D_{\mathbf{hw}}^{*}\right)\cdot \\ &= \rho(1-\rho)^3 p\left(d_{4\to3} < D_{4\to3}^{RSS}\right)\cdot p\left(d_{3\to2} < D_{\mathbf{hw}}^{*}\right) \end{aligned} \tag{22}$$

$D_{\mathbf{hw}}^{*}$ is the critical dilemma distance specified by (9) and (10). The probability that the trilemma occurs while there is no dilemma of the leading three vehicles is calculated as:

$$\begin{aligned} \mathrm{p_{trilemma\backslash\,dilemma}} &= p(HHAH)\cdot p\left(d_{4\to3} < D_{4\to3}^{RSS}\right) \\ &\quad \cdot p\left(D_{3\to2}^{RSS} < d_{3\to2} < D_{\mathbf{hw}}^{*}\right) \\ &= \rho(1-\rho)^3 p\left(d_{4\to3} < D_{4\to3}^{RSS}\right) \\ &\quad \cdot p\left(D_{3\to2}^{RSS} < d_{3\to2} < D_{\mathbf{hw}}^{*}\right) \end{aligned} \tag{23}$$

Therefore, the risk reduction probability when one AV is presented $p_{reduce|1\ AV}$ can be expressed as:

$$\begin{aligned} p_{reduce|1\ AV} &= p_{\mathrm{D}\,ilemma} + p_{\mathrm{trilemma\backslash\,dilemma}} \\ &= \rho(1-\rho)^2\cdot p\left(d_{3\to2} < D_{3\to2}^{RSS}\right) \\ &\quad +\rho(1-\rho)^3 p\left(d_{4\to3} < D_{4\to3}^{RSS}\right) \\ &\quad \cdot p\left(D_{3\to2}^{RSS} < d_{3\to2} < D_{\mathbf{hw}}^{*}\right) \\ &= \rho(1-\rho)^2 F_{\mathrm{hw}}\left(D_{3\to2}^{RSS}|v\right) + \rho(1-\rho)^3 \\ &\quad \cdot F_{\mathrm{hw}}\left(D_{4\to3}^{RSS}|v\right)\left(F_{\mathrm{hw}}\left(D_{\mathbf{hw}}^{*}|v\right) - F_{\mathrm{hw}}\left(D_{3\to2}^{RSS}|v\right)\right) \end{aligned} \tag{24}$$

Next we consider the case when there are arbitrary N AVs at the middle, i.e., $\mathrm{H}\overbrace{A\cdots A}^{N}\mathrm{H}$ and $\mathrm{HH}\overbrace{A\cdots A}^{N}\mathrm{H}$, where $\mathrm{N} \geq 2$. The occurrence probability of the platoon can be expressed as follows:

$$\begin{cases} \mathrm{p}\left(\mathrm{H}\overbrace{A\cdots A}^{N}\mathrm{H}\right) = (1-\rho)\rho^N(1-\rho) = \rho^N(1-\rho)^2 & (a) \\ \mathrm{p}\left(\mathrm{HH}\overbrace{A\cdots A}^{N}\mathrm{H}\right) = (1-\rho)(1-\rho)\rho^N(1-\rho) = \rho^N(1-\rho)^3 & (b) \end{cases} \tag{25}$$

Then, the probability of $\mathrm{N}+1$*Lemma* and $\mathrm{N}+2$*Lemma* are calculated as follows:

$$\mathrm{p_{N+1}}\mathit{Lemma} = \mathrm{p}\left(\mathrm{H}\overbrace{A\cdots A}^{N}\mathrm{H}\right)\cdot \mathrm{p}\left(d_{N+2\to N+1} < D_{N+2\to N+1}^{RSS}\right)$$

$$= \rho^N(1-\rho)^2 \mathrm{F_{hw}}\left(D^{RSS}_{N+2\to N+1}\right) \tag{26}$$

The probability of N + 2*Lemma* is calculated as follows

$$\begin{aligned}
&\mathrm{p_{N+2Lemma}} = \mathrm{p}\left(\mathrm{HH}\overbrace{A\cdots A}^{N}\mathrm{H}\right)\\
&\cdot \mathrm{p}\left(d_{N+3\to N+2} < D^{RSS}_{N+3\to N+2}\right) p\left(d_{N+2\to N+1} < D^{*}_{\mathbf{hw}}\right)\\
&= \rho^N(1-\rho)^3 \mathrm{F_{hw}}\left(D^{RSS}_{N+3\to N+2}\right) F_{\mathrm{hw}}\left(D^{*}_{\mathbf{hw}}\right)
\end{aligned} \tag{27}$$

The probability of N+2*Lemma* occurrence which exclude the N + 1*Lemma* event is expressed as:

$$\begin{aligned}
&\mathrm{p_{N+2Lemma\backslash N+1Lemma}}\\
&= \mathrm{p}\left(\mathrm{HH}\overbrace{A\cdots A}^{N}\mathrm{H}\right)\\
&\cdot \mathrm{p}\left(d_{N+3\to N+2} < D^{RSS}_{N+3\to N+2}\right)\\
&\cdot \mathrm{p}\left(D^{RSS}_{N+3\to N+2} < h_{N+2\to N+1} < D^{*}_{\mathbf{hw}}\right)\\
&= \rho^N(1-\rho)^3 F_{\mathrm{hw}}\left(D^{RSS}_{N+3\to N+2}\right)\\
&\left[\mathrm{F_{hw}}\left(D^{*}_{\mathbf{hw}}\right) - \mathrm{F_{hw}}\left(D^{RSS}_{N+3\to N+2}\right)\right]
\end{aligned}$$

Hence, when there are N AVs, the risk reduction can be calculated as

$$\begin{aligned}
\mathrm{p}_{reduce|N\ AVs} &= \mathrm{P}_{N+1Lemma} + \mathrm{P}_{N+2Lemma\backslash N+1Lemma}\\
&= \rho^N(1-\rho)^2\,\mathrm{F_{hw}}\left(D^{RSS}_{N+2\to N+1}\right)\\
&\quad + \rho^N(1-\rho)^3\,\mathrm{F_{hw}}\left(D^{RSS}_{N+3\to N+2}\right)\\
&\quad \left[\mathrm{F_{hw}}\left(D^{*}_{\mathbf{hw}}\right) - F_{\mathrm{hw}}\left(D^{RSS}_{N+3\to N+2}\right)\right]
\end{aligned} \tag{28}$$

Finally, the total risk reduction when the MPR is $\rho$ for the speed v is given by:

$$\begin{aligned}
\mathrm{p}_{reduce} &= \sum_{N=1}^{\infty} \mathrm{p}_{reduce|N\ AVs}\\
&= \sum_{N=1}^{\infty} \rho^N(1-\rho)^2 \mathrm{F_{hw}}\left(\mathrm{D}^{RSS}\right) + \rho^N(1-\rho)^3\\
&\quad \cdot \mathrm{F_{hw}}\left(\mathrm{D}^{RSS}\right)\left(\mathrm{F_{hw}}\left(D^{*}_{\mathbf{hw}}\right) - \mathrm{F_{hw}}\left(\mathrm{D}^{RSS}\right)\right)
\end{aligned} \tag{29}$$

### C. EFFICIENCY COMPROMISING INVESTIGATION VIA THE CAPACITY

#### 1) AVERAGE SPACE GAP DISTRIBUTION OF A GIVEN POLYLEMMA PLATOON

We first consider the distribution of average space gap $\bar{h}_{\mathrm{dilemma}}$ for critical dilemma scenario in platoon Fig. 1-a. The average space gap $\bar{h}_{\mathrm{dilemma}}$ is $\frac{2l_{\mathrm{veh}}+d_{3\to 2}+D^{*}_{\mathbf{hw}}}{2} - \mathrm{l_{veh}} = \frac{\mathrm{d_{3\to 2}}+\mathrm{D^{*}_{hw}}}{2}$. Because $d_{3\to 2} \in (0, D^{RSS}_{3\to 2})$, $\bar{h}_{\mathrm{dilemma}}$ then falls within the interval of $(\frac{D^{*}_{\mathbf{hw}}}{2}, \frac{D^{RSS}_{3\to 2}+D^{*}_{\mathbf{hw}}}{2})$. Then, the distribution of $\bar{h}_{\mathrm{dilemma}}$ given v and $\rho$ can be calculated as follows:

$$\begin{aligned}
&p\left(\bar{h}_{\mathrm{dilemma}} = h|v,\rho,1AV\right) = \mathrm{p}\left(\frac{d_{3\to 2}+D^{*}_{\mathbf{hw}}}{2} = h|v\right)\\
&= \mathrm{p}\left(d_{3\to 2} = 2h - D^{*}_{\mathbf{hw}}|v\right)\\
&= \begin{cases} \mathrm{f_{hw}}\left(2h - D^{*}_{\mathbf{hw}}|v\right), \forall h \in \left(\frac{D^{*}_{\mathbf{hw}}}{2}, \frac{D^{RSS}_{3\to 2}+D^{*}_{\mathbf{hw}}}{2}\right) \\ 0, \mathit{else} \end{cases}
\end{aligned} \tag{30}$$

In a similar derivation, we calculate the headway distribution for trilemma in scenario Fig. 3-a. The average headway is $\frac{3l_{\mathrm{veh}}+d_{4\to 3}+d_{3\to 2}+D^{*}_{\mathbf{hw}}}{3} - \mathrm{l_{veh}} = \frac{d_{4\to 3}+d_{3\to 2}+D^{*}_{\mathbf{hw}}}{3}$. Then:

$$\begin{aligned}
&p\left(\bar{h}_{trilemma\backslash\,\mathrm{dilemma}} = h|v,\rho,1AV\right)\\
&= \iint p\left(\bar{h}_{trilemma\backslash\,\mathrm{dilemma}} = h|d_{3\to 2}, d_{4\to 3}\right) p(d_{3\to 2}, d_{4\to 3}|v)\,d\,d_{3\to 2}\,d\,d_{4\to 3}\\
&= \int_0^{D^{RSS}_{4\to 3}} \left[\int_{D^{RSS}_{3\to 2}}^{D^{*}_{\mathbf{hw}}} p\left(\frac{d_{4\to 3}+d_{3\to 2}+D^{*}_{\mathbf{hw}}}{3} = h|d_{4\to 3}, d_{3\to 2}\right) p(d_{4\to 3}, d_{3\to 2}|v)\,d\,d_{3\to 2}\right]\\
&\quad d\,\mathrm{D}_{4\to 3}\\
&= \int_0^{D^{RSS}_{4\to 3}} \left[\int_{D^{RSS}_{3\to 2}}^{D^{*}_{\mathbf{hw}}} p\left(d_{4\to 3}+d_{3\to 2} = 3h - D^{*}_{\mathbf{hw}}|d_{4\to 3}+d_{3\to 2}\right) f_{\mathrm{hw}}(d_{4\to 3}) f_{\mathrm{hw}}\right.\\
&\quad \left.(d_{3\to 2})dx_{3\to 2}\right] d\ \mathrm{D}_{4\to 3}
\end{aligned} \tag{31}$$

In the above formulation, the probability $p(d_{4\to 3}+d_{3\to 2} = vh - D^{*}_{\mathbf{hw}}|d_{4\to 3}, d_{3\to 2})$ can be expressed using Dirac function, i.e.:

$$\begin{aligned}
&p\left(d_{4\to 3}+d_{3\to 2} = 3h - D^{*}_{\mathbf{hw}}|d_{4\to 3}, d_{3\to 2}\right)\\
&= Dirac\left(d_{4\to 3}+d_{3\to 2}, 3h - D^{*}_{\mathbf{hw}}\right)\\
&= \begin{cases} 1, \ \mathit{if}\ d_{4\to 3}+d_{3\to 2} = 3h - D^{*}_{\mathbf{hw}} \\ 0, \mathit{else} \end{cases}
\end{aligned} \tag{32}$$

Therefore:

$$\begin{aligned}
&p\left(\bar{h}_{trilemma\backslash\,\mathrm{dilemma}} = h|v,\rho,1AV\right)\\
&= \int_0^{D^{RSS}_{4\to 3}} \left[\int_{D^{RSS}_{3\to 2}}^{D^{*}_{\mathbf{hw}}(x_{34})} \mathrm{Dirac}\left(\mathrm{d_{4\to 3}}+\mathrm{d_{3\to 2}}, 3\mathrm{h} - \mathrm{D^{*}_{hw}}\right)\right.\\
&\quad \left.\mathrm{f_{hw}}(d_{4\to 3})\mathrm{f_{hw}}(d_{3\to 2})d\ d_{3\to 2}\right] d\ d_{4\to 3}
\end{aligned} \tag{33}$$

The probability $\mathrm{p}(\bar{\mathrm{h}}_{\mathrm{dilemma}} = \mathrm{h}|\mathrm{v},\ 1\ \mathrm{AV})$ and $p(\bar{h}_{trilemma\backslash\,\mathrm{dilemma}} = h|v, 1AV)$ correspond to the scenario when there is only one AV, followed by other HDVs. The case can be generalized to arbitrary number (denoted as N) of AVs, represented by $p(\bar{h}_{N+1Lemma} = h|v, N\ AVs)$ and $p(\bar{h}_{N+2Lemma\backslash N+1Lemma} = h|v, N\ AVs)$. The distribution of the average headway for a platoon of vehicles under the critical polylemma scenario is mathematically expressed as follows:

$$\begin{aligned}
&p\left(\bar{h}_{\mathrm{polylemma}}|v,\rho\right) = \sum_{N=2}^{\infty} \rho^N(1-\rho)^3\\
&\cdot \mathrm{p}\left(\bar{h}_{N+2Lemma\backslash N+1Lemma} = h|v, N\ AVs\right)\\
&+\rho^N(1-\rho)^2 \cdot \mathrm{p}\left(\bar{h}_{N+1Lemma} = h|v, N\ AVs\right)
\end{aligned} \tag{34}$$

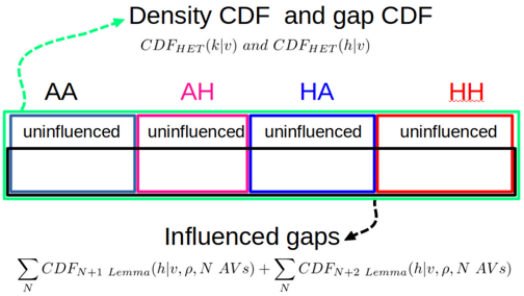


**FIGURE 7.** Event space decomposition.

It is shown that the above formulation is not explicitly expressed and cannot be solved directly. We have to resort to numerical methods as described in the following section.

### 2) NUMERICAL SCHEME TO OBTAIN THE DISTRIBUTION OF

The event space for pairs of vehicles in a mixed flow of autonomous vehicles (AVs) consists of four types: AA, AH, HA, and HH. These correspond to the probabilities $\rho^2$, $\rho(1-\rho)$, $\rho(1-\rho)$ and $(1-\rho)^2$. Fig. 7 depicts the decomposition of this event space, with gaps classified into two categories: those influenced by the HET (lower part) and those unaffected (upper part). The gaps for AA and AH, which are not influenced by the HET, are given by $\frac{1}{k^{eq}_{AA}(v)} - l_{\text{veh}}$ and $\frac{1}{k^{eq}_{AH}(v)} - l_{\text{veh}}$ and are deterministic. To alleviate computational complexity, the Cumulative Distribution Function is employed. Specifically, the cumulative distribution of the unaffected gaps for AA, AH, HA, and HH can be denoted as $\mathcal{S}(h \dashv \frac{1}{k^{eq}_{AA}(v)} - l_{\text{veh}})$ and $\mathcal{S}(h \dashv \frac{1}{k^{eq}_{AH}(v)} - l_{\text{veh}})$. $F_{\text{hw}}(h|v)$, $F_{\text{hw}}(h|v)$ respectively. The function $\mathcal{S}(h \dashv h^{\#})$ is used to express the CDF and is defined as:

$$\mathcal{S}(h \dashv \hbar) \overset{\text{Def}}{=} \begin{cases} 0 & \forall h \in (-\infty, \hbar) \\ 1 & \forall h \in [\hbar, \infty] \end{cases} \tag{35}$$

$F_{\text{hw}}(h|v)$ is the gap distribution derived from speed-density relationship for certain background traffic flow as in Fig. 6-b. In summary, the cumulative distribution of gap h, without the application of HET, can be calculated as follows by decomposing the event space into four types as in Fig. 7, considering a given MPR $\rho$ and speed v:

$$\begin{aligned} \text{CDF}(h|v,\rho) = {} & \rho^2 \cdot \mathcal{S}\left(h \dashv \frac{1}{k^{eq}_{AA}(v)} - l_{\text{veh}}\right) \\ & + \rho(1-\rho) \cdot \mathcal{S}\left(h \dashv \frac{1}{k^{eq}_{AH}(v)} - l_{\text{veh}}\right) \\ & + \rho(1-\rho) \cdot F_{\text{hw}}(h|v) + (1-\rho)^2 F_{\text{hw}}(h|v) \end{aligned} \tag{36}$$

The cumulative distribution of density and gap after the application of polylemma avoidance are denoted as $\text{CDF}_{\text{HET}}(k|v,\rho)$ and $\text{CDF}_{\text{HET}}(h|v,\rho)$. The calculation of $\text{CDF}_{\text{HET}}(h|v,\rho)$ can be calculated by combining the average gaps that are altered and those not:

Step 1. For each N (number of consecutive AVs) ranging from 1 to infinity, we compute the cumulative probability of the average space gap $h'$ when N+1Lemma avoidance (for platoon structure $\text{H}\overbrace{A\cdots A}^{N}H$) and N+2*Lemma* avoidance (for platoon structure $\text{HH}\overbrace{A\cdots A}^{N}H$) are implemented, which are denoted as $CDF_{N+1Lemma}(h'|v,\rho,N\ AVs)$ and $CDF_{N+2\ Lemma}(h'|v,\rho,N\ AVs)$ respectively;

Step 2. In the platoon scenario $H\overbrace{A\cdots A}^{N}H$, the number of occurrence for HA, AA and AH are 1, N-1 and 1 respectively. Therefore the occurrence probabilities of the above three combinations which are influenced by polylemma avoidance are $(1-\rho)^2\rho^N$, $(1-\rho)^2\rho^N(N-1)$ and $(1-\rho)^2\rho^N$;

Step 3. Similarly, in the platoon scenario $\text{HH}\overbrace{A\cdots A}^{N}H$, the number of combinations for HH, HA, AA and AH are 1, 1, N-1 and 1 respectively. Therefore the occurrence probabilities of the above four combinations which are influenced by polylemma avoidance in N-AVs platoon are $(1-\rho)^3\rho^N$, $(1-\rho)^3\rho^N$, $(1-\rho)^3\rho^N(N-1)$ and $(1-\rho)^3\rho^N$ respectively.

The occurrence probabilities of the unaffected AA, AH, HA, and HH pairs are obtained by subtracting the influenced probabilities from the original probabilities:

$$\begin{cases} \textit{uninfluenced AA prob}\text{: } \rho^2 - (1-\rho)^2\rho^N(N-1) - (1-\rho)^3\rho^N(N-1) & (a) \\ \textit{uninfluenced AH prob}\text{: } \rho(1-\rho) - (1-\rho)^2\rho^N - (1-\rho)^3\rho^N & (b) \\ \textit{uninfluenced HA prob}\text{: } \rho(1-\rho) - (1-\rho)^2\rho^N - (1-\rho)^3\rho^N & (c) \\ \textit{uninfluenced HH prob}\text{: } (1-\rho)^2 - (1-\rho)^3\rho^N & (d) \end{cases} \tag{37}$$

Finally, the $\text{CDF}_{\text{HET}}(h|v,\rho)$ can be obtained via adding the uninfluenced gaps with the influenced counterparts, which include $\sum_N CDF_{N+1Lemma}(h|v,\rho,N\ AVs)$ and $\sum_N CDF_{N+2Lemma}(h|v,\rho,N\ AVs)$.

$$\begin{aligned} & \text{CDF}_{\text{HET}}(h|v,\rho) \\ & = \sum_N \left[\rho^2 - (1-\rho)^2\rho^N(N-1) - (1-\rho)^3\rho^N(N-1)\right] \\ & \cdot \mathcal{S}\left(h \dashv \frac{1}{k^{eq}_{AA}(v)} - l_{\text{veh}}\right) \\ & + \sum_N \left[\rho(1-\rho) - (1-\rho)^2\rho^N - (1-\rho)^3\rho^N\right] \\ & \cdot \mathcal{S}\left(h \dashv \frac{1}{k^{eq}_{AH}(v)} - l_{\text{veh}}\right) \\ & + \sum_N \left[\rho(1-\rho) - (1-\rho)^2\rho^N - (1-\rho)^3\rho^N\right] \\ & \cdot F_{\text{hw}}(h + l_{\text{veh}}|v) \\ & + \sum_N \left[(1-\rho)^2 - (1-\rho)^3\rho^N\right] F_{\text{hw}}(h + l_{\text{veh}}|v) \end{aligned}$$

$$+\left(\sum_{N} CDF_{N+1Lemma}(h|v,\rho,N\ AVs)\right)$$
$$+\left(\sum_{N} CDF_{N+2\ Lemma}(h|v,\rho,N\ AVs)\right) \quad (38)$$

The unknown terms in (38) include $CDF_{N+1Lemma}(h|v,\rho,N\ AVs)$ and $CDF_{N+2\ Lemma}(h|v,\rho,N\ AVs)$. We now turn to the discussion of the calculation of the above two CDFs given that there are N AVs in the platoon. The calculation steps of the $CDF_{N+1Lemma}(h|v,\rho,N\ AVs)$ and $CDF_{N+2\ Lemma}(h|v,\rho,N\ AVs)$ are developed as follows.

1) We first Consider the case of N+1*Lemma*, i.e., the platoon structure is $H\overbrace{A\cdots A}^{N}H$, which occurs with probability $(1-\rho)^2\rho^N$. The gap between the last HDV and second to the last AV is h and its distribution is $f_{\text{hw}}(h|v)$, h$\in [0, D^{RSS}(v)]$;
2) The average headway $h'$ after HET is applied is calculated as following:

$$h' = \frac{h+(N-1)\cdot\left(\frac{1}{k_{av}(v)}-l_{\text{veh}}\right)+\mathfrak{D}_{\text{hw}}^{N+1}+(N+1)\cdot l_{\text{veh}}}{N+1}$$
$$-l_{\text{veh}} = \frac{h+(N-1)\cdot\left(\frac{1}{k_{av}(v)}-l_{\text{veh}}\right)+\mathfrak{D}_{\text{hw}}^{N+1}}{N+1} \quad (39)$$

The terms in the numerator of the first term after the first equals sign in (39) represent the gap of the last HDV, the gap of the last N-1 AVs, the gap of the first AV (denoted as $\mathfrak{D}_{\text{hw}}^{N+1}$ to avoid the polylemma), and the length of the vehicles. By rearranging (39), we can express h in terms of h', i.e., h$= (N+1)h' - \mathfrak{D}_{\text{hw}}^{N+1} - (N-1)\cdot(\frac{1}{k_{av}(v)} - l_{\text{veh}})$. Thus, $CDF_{N+1Lemma}(h'|v,\rho,N\ AVs)$ can be expressed as follows:

$$CDF_{N+1Lemma}\left(h'|v,\rho,N\ AVs\right)$$
$$= F_{\text{hw}}\left((N+1)h' - \mathfrak{D}_{\text{hw}}^{N+1} - (N-1)\cdot\left(\frac{1}{k_{av}(v)}-l_{\text{veh}}\right)|v\right) \quad (40)$$

3) Next, we consider N+2*Lemma* with platoon structure $HH\overbrace{A\cdots A}^{N}H$ with occurrence probability $(1-\rho)^3\rho^N$. The necessary conditions of N+2*Lemma* given the platoon structure are that the gap of the last HDV, i.e., $h$, satisfies $0 < h < D^{RSS}(v)$ while the gap of the second to last HDV, i.e., $h''$, satisfies $h'' < D^{*}_{\mathbf{hw}}(h,v)$. The average headway $h'$ is calculated as:

$$h' = \frac{h+h''+(N-1)\cdot\left(\frac{1}{k_{av}(v)}-l_{\text{veh}}\right)+\mathfrak{D}_{\text{hw}}^{N+2}(v)+(N+2)\cdot l_{\text{veh}}}{N+2}$$
$$-l_{\text{veh}} \quad (41)$$

The structure of (41) is similar as (39). Given a fixed h, the average gap $h'$ can be used to express $h''$ by rearranging (41):

$$h'' = (N+2)h' - (N-1)\cdot\left(\frac{1}{k_{av}(v)}-l_{\text{veh}}\right)$$
$$+\mathfrak{D}_{\text{hw}}^{N+2}(v) - h \quad (42)$$

The distribution $h''$ is exactly the gap distribution $f_{\text{hw}}(h''|v)$. Hence distribution of $h'$, which is indicated using $p_{N+2\ Lemma}(h'|h,v,\rho,N\ AVs)$ is expressed using $f_{\text{hw}}(h''|v)$ as follows:

$$\begin{aligned}
&p_{N+2\ Lemma}\left(h'|h,v,\rho,N\ AVs\right)\\
&= \text{p}\Big(h'' = (N+2)h' - (N-1)\cdot\left(\frac{1}{k_{av}(v)}-l_{\text{veh}}\right) + \mathfrak{D}_{\text{hw}}^{N+2}(v) + (N+2)\cdot l_{\text{veh}}\\
&\qquad -h|h,v,\rho,N\ AVs\Big)\\
&= f_{\text{hw}}\left((N+2)h' - (N-1)\cdot\left(\frac{1}{k_{av}(v)}-l_{\text{veh}}\right) + \mathfrak{D}_{\text{hw}}^{N+2}(v) + (N+2)\cdot l_{\text{veh}} - h|v\right)
\end{aligned} \quad (43)$$

To eliminate h in $p_{N+2\ Lemma}(h'|h,v,\rho,N\ \ AVs)$, we integrate out h, which is expressed as:

$$\begin{aligned}
&p_{N+2\ Lemma}\left(h'|v,\rho,N\ AVs\right)\\
&= \int_0^{D^{RSS}(v)} p_{N+2\ Lemma}\left(h'|h,v,\rho,N\ AVs\right)dh\\
&= \int_0^{D^{RSS}(v)} f_{\text{hw}}\left((N+2)h' - (N-1)\cdot\left(\frac{1}{k_{av}(v)}-l_{\text{veh}}\right) + \mathfrak{D}_{\text{hw}}^{N+2}(v) - h|v\right)dh\\
&= \sum_i \Delta h\cdot f_{\text{hw}}\left((N+2)h' - (N-1)\cdot\left(\frac{1}{k_{av}(v)}-l_{\text{veh}}\right) + \mathfrak{D}_{\text{hw}}^{N+2}(v) - h_i|v\right)
\end{aligned} \quad (44)$$

In the final derivation outlined in (39), the integration of probability is solved using numeral summation. The CDF $\text{CDF}_{N+2Lemma}(h'|v,\rho,N\ AVs)$ represents the integral of the PDF, which is expressed as:

$$\begin{aligned}
&\text{CDF}_{N+2Lemma}\left(h'|v,\rho,N\ AVs\right)\\
&= \int_0^{h'} p_{N+2\ Lemma}\left(h'|v,\rho,N\ AVs\right)d\mathcal{H}\\
&= \int_0^{h'}\left[\int_0^{D^{RSS}(v)} f_{\text{hw}}\left((N+2)h' - (N-1)\cdot\left(\frac{1}{k_{av}(v)}-l_{\text{veh}}\right) + \mathfrak{D}_{\text{hw}}^{N+2}(v) - \mathcal{H}|v\right)dh\right]d\mathcal{H}\\
&= \sum_j \Delta\mathcal{H}\left[\sum_i \Delta h\cdot f_{\text{hw}}\left((N+2)\mathcal{H}_j - (N-1)\cdot\left(\frac{1}{k_{av}(v)}-l_{\text{veh}}\right) + \mathfrak{D}_{\text{hw}}^{N+2}(v) - h_i|v\right)\right]
\end{aligned} \quad (45)$$

As the density k is computed through h, i.e., $\text{k} = \frac{1}{h+l_{\text{veh}}}$, the $\text{CDF}_{\text{HET}}(k|v,\rho)$ can thus obtained.

## VI. CASE STUDY

### A. DATA SOURCES

The dataset of this subsection comes from NGSIM. The NGSIM data is a popular trajectory data that contain vehicle movements in multi lane road. Because the focus of this research is longitudinal driving, the platoon data is extracted from the raw dataset. They will be used to examine the

### B. DILEMMA AND TRILEMMA IN EXAMPLE PLATOON

For the dilemma scenario, we select a platoon consisting of three vehicles, with the leader labeled as vehicle 0, and the second and third vehicles labeled as 1 and 2, respectively (Fig. 8). The time horizon of the platoon is approximately 60 seconds. The speed profiles and longitudinal coordinates of the three vehicles are depicted in Fig. 8 (left sub-figure

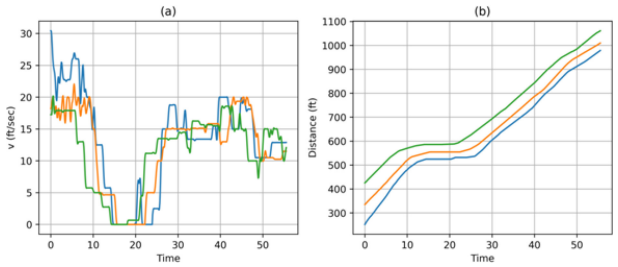


**FIGURE 8.** Speed profile and longitudinal location for 3-vehicles platoon.

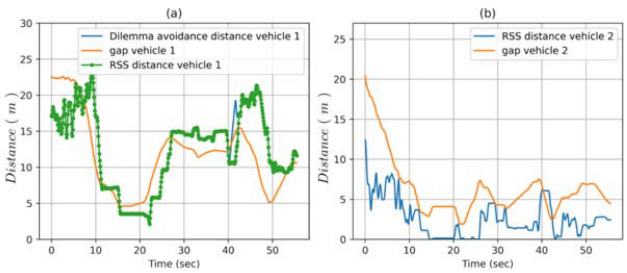


**FIGURE 9.** Gap, RSS distance and Dilemma avoidance distance.

for speed profiles and right sub-figure for longitudinal coordinates). To compute the minimum safe distance (RSS distance), we employ (1) and (2), which require parameters such as the deceleration of the two vehicles and the reaction delay. The response time is estimated using the cross-correlation method described in [55]. The lag is determined for each vehicle, and subsequently, the RSS distance is calculated.

Fig. 9 presents the values for the RSS distance, space gap, and dilemma avoidance distance in a typical platoon scenario. Specifically, Fig. 9-a illustrates these distances for second vehicle (vehicle 1) in the platoon, while Fig. 9-b portrays them for the last vehicle. In Fig. 9-b, it is evident that the gap exceeds the RSS distance for the majority of the time, with the exception occurring between 40.3 sec and 42.3 sec. This discrepancy is due to the slight deceleration of the middle vehicle, which reduces the gap as indicated in Fig. 8-b. Consequently, the dilemma scenario arises due to the violation of the RSS rule, necessitating a larger distance known as the dilemma avoidance distance. Fig. 9-a further depicts this occurrence between 40.3 sec and 42.3 sec, where the magnitude of the dilemma avoidance distance surpasses the corresponding gap. The duration of RSS violation is found to be 59.3% for vehicle 1 and 3.95% for vehicle 2. Notably, as the RSS violation of vehicle 2 leads to the dilemma condition, the dilemma occurrence ratio for this platoon dataset also stands at 3.95%.

Next we investigate the trilemma in real world data. This analysis focuses on a four-vehicle platoon, specifically examining the occurrence of a trilemma situation. Fig. 10 showcases the speed profiles and longitudinal coordinates of the platoon. The platoon encounters a stop-and-go event from approximately 80 seconds to 100 seconds (observed from the speed in Fig. 10-a). Fig. 11 presents the RSS distance,

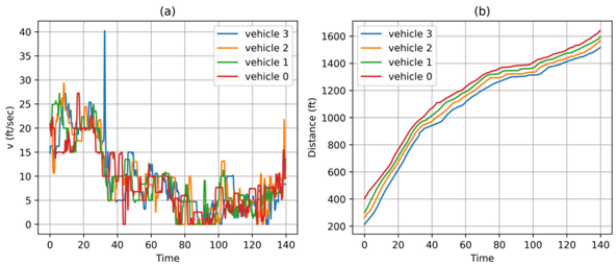


**FIGURE 10.** Speed profiles and longitudinal locations for a 4-vehicle platoon.

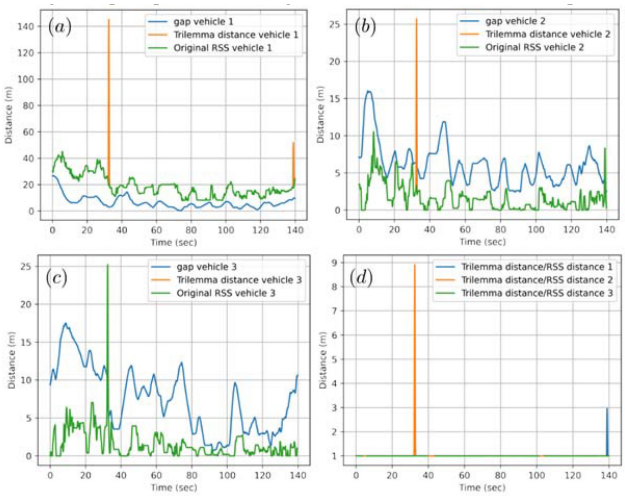


**FIGURE 11.** Gap distance, trilemma distance and RSS distance.

gap, and trilemma distance for the platoon, with sub-figures (a), (b), and (c) showcasing these distances individually. For vehicle 3 (depicted in Fig. 11-c), the majority of the time exhibits a safe situation, with the RSS distance smaller than the gap. However, within a brief interval of 0.7 seconds, specifically from 32.2 seconds to 32.9 seconds, RSS distance significantly increases. This occurrence stems from the acceleration of vehicle 3, evident in Fig. 10-a, and the resulting reduction in the corresponding gap, as depicted in Fig. 10-b. As the RSS distance for vehicle 3 increases, it requires vehicle 2 to maintain a larger gap with its leader, illustrated in Fig. 11-b, corresponding to a distance exceeding 25 meters. Consequently, this necessitates vehicle 1 to maintain an even larger gap, exceeding 140 meters as depicted in Fig. 11-a. The duration of the trilemma occurrence ratio accounts for approximately 0.5% of the entire time horizon, specifically 0.7 seconds within 140 seconds. Fig. 11-d portrays the ratio of the trilemma distance to the RSS distance, with values exceeding 1 indicating a trilemma situation. The maximum ratio observed is approximately 9 during the interval from 32.2 seconds to 32.9 seconds, indicating that vehicle 1 needs to expand its gaps to nine times their original size.

### C. RATIOS OF RSS VIOLATION AND POLYLEMMA IN REAL WORLD DATASET

We examine two publicly available datasets: highD and zen-dataset, which represent traffic data from Germany and

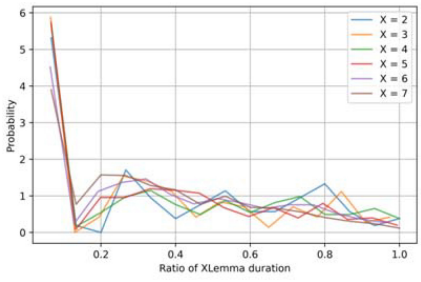


**FIGURE 12.** Distributions of RSS violation, dilemma and trilemma for highD dataset.

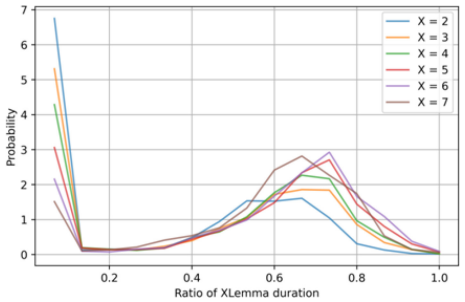


**FIGURE 13.** Distributions of RSS violation, dilemma and trilemma for zen dataset.

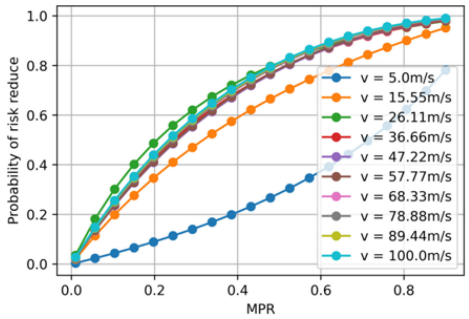


**FIGURE 14.** The probability of risk reduction.

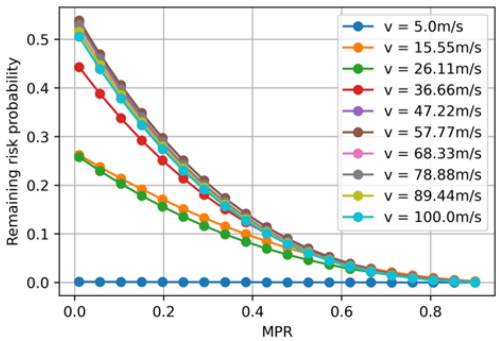


**FIGURE 15.** The probability of remaining risk under different MPR.

Japan, respectively. Given a platoon with a duration T, we calculate the occurrence duration of polylemma. To illustrate the violation ratio of the dilemma in traffic flow, we use the example of a platoon consisting of three vehicles labeled 1, 2, and 3, with vehicle 1 being the leader. The platoon data is considered within the time horizon [0,T]. The reaction times of vehicle 2 and 3 are denoted as $\tau_2$ and $\tau_3$ respectively. For each moment t $\in [\tau_2, T]$, we first compute the required RSS distance $D^{RSS}_{3\to 2}$ at that moment. If the distance $d_{3\to 2}(t)$ is smaller than $D^{RSS}_{3\to 2}$, it indicates a violation by vehicle 3. Next, we calculate the RSS distance $D^{RSS}_{2\to 1}$ and the critical dilemma distance $D^{*}_{hw}$ for vehicle 2. If the condition $D^{RSS}_{2\to 1} < d_{2\to 1}(t) < D^{*}_{hw}$ and $d_{3\to 2}(t) < D^{RSS}_{3\to 2}$ are simultaneously met, it confirms the occurrence of a dilemma instance. Let n denote the number of dilemma instances during the platoon, then the ratio of dilemma violations is given by $\eta_{\mathrm{D}\,ilemma} = \frac{n}{N}$ for this platoon, where N is the total instants of the data. The calculation of polylemma occurrences follows a similar process.

The distribution of the polylemma ratio in each dataset exhibits specific characteristics. In Fig. 12, histograms representing the ratios of polylemma instances in the highD dataset are displayed for different values of X. Each histogram curve is normalized, ensuring that the integral of each curve is equal to 1. These histograms reveal that the occurrence of polylemma is not prevalent in most circumstances, as indicated by the peak probability at a ratio of 0. However, there are instances within the platoon data where polylemma occurs consistently, as evidenced by positive distribution values associated with a ratio of 1. When polylemma is present, the most probable ratio is approximately 30%.

Fig. 13 showcases the distribution of polylemma occurrence ratios in the zen traffic dataset. Contrasting with the findings in the highD dataset, zen traffic data contains a larger amount of data resulting in a smoother curve (also normalized to have an integral of 1), and several differences emerge. Firstly, it is evident that for each value of X, a significant ratio of no-XLemma platoons can be observed. However, this ratio diminishes as X increases. Moreover, for a given X value, the most probable ratio, if not 0, varies. In the case of dilemma, the most probable ratio is 60%, whereas for hexalemma (X=6), the most probable ratio surpasses 70%. These findings imply that the risk associated with platoons in the zen dataset is greater than that in the highD dataset.

### D. RISK REDUCTION

The risk reduction intuitively gives the safety potential of the proposed HET strategy. We take the traffic flow in NGSIM (i.e., Fig. 6) as the background traffic flow. As aforementioned, we increase the MPRs of the background traffic flow and assume that the remaining HDVs behaviors (gap distribution) are not changed. The risk reduction rate and remaining risk probability is given in Fig. 14 and Fig. 15 for different speeds. Fig. 14 is the results of risk reductio. When MPR is 100%, all risks are eliminated. The risk reduction increases with the MPR, but varies across different speeds, which is due to the fact that the RSS violation is different under different speed in Fig. 6. Fig. 15 presents the remaining risk after the polylemma avoidance strategy is

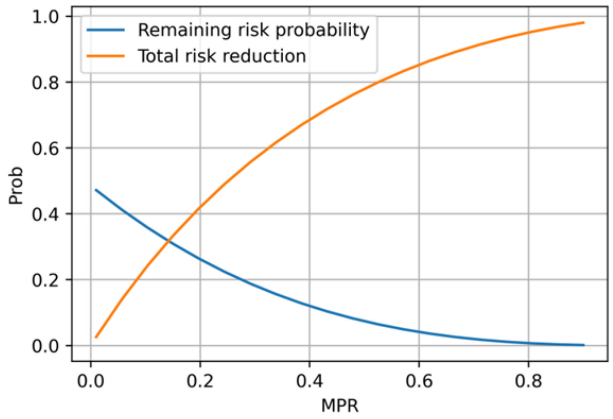


**FIGURE 16.** Total risk reduction and remaining risk under different MPR.

applied. When the MPR is not 100%, there are always certain HDVs that violate the RSS rule and thus still there would be certain risk probability. The risk generally decreases against the MPR. At v=57.77m/s, the risk occurrence probability is about 13% for the case when MPR is 50%. The total risk reduction for the whole traffic flow under fixed MPR can be calculated by weighted integration considering the speed distribution in Fig. 6, which are given in Fig. 16. It is shown that 50% MPR would reduce over 80% risks due to human drivers' error, while the corresponding occurrence probability of remaining risk in the traffic flow is only about 10%, compared with about 50% of zero MPR.

### *E. SPEED-DENSITY RELATIONSHIP AND CAPACITY INFLUENCE OF HET*

We still use the NGSIM in Fig. 6 as background traffic flow. Fig. 6-b is the raw speed-density relationship. As aforementioned, we use the CDF function $\mathrm{CDF}_{\mathrm{HET}}(k|v)$ to calculate the conditional probability of density, and the corresponding expectation of density can be derived using the following numerical formula by employing $\mathrm{CDF}_{\mathrm{HET}}(k|v)$:

$$E(k|v) = \int k \cdot PDF_{\mathrm{HET}}(k|v)dk = \int k \cdot \frac{d\ \mathrm{CDF}_{\mathrm{HET}}(k|v)}{dk} dk$$
$$= \sum_{\mathrm{i}} \mathrm{k_i} \left.\frac{\mathrm{d}\ \mathrm{CDF}_{\mathrm{HET}}(\mathrm{k}|\mathrm{v})}{\mathrm{dk}}\right|_{\mathrm{k_i}} \Delta\mathrm{k}$$
$$= \sum_{\mathrm{i}} \mathrm{k_i} \left.\frac{(\mathrm{CDF}_{\mathrm{HET}}(\mathrm{k_{i+1}}|\mathrm{v}) - \mathrm{CDF}_{\mathrm{HET}}(\mathrm{k_i}|\mathrm{v}))}{\Delta\mathrm{k}}\right|_{\mathrm{k_i}} \Delta\mathrm{k}$$
$$= \sum_{\mathrm{i}} \mathrm{k_i}((\mathrm{CDF}_{\mathrm{HET}}(\mathrm{k_{i+1}}|\mathrm{v}) - \mathrm{CDF}_{\mathrm{HET}}(\mathrm{k_i}|\mathrm{v}))) \quad (46)$$

In above equation, $\{k_i\}$, $k_i \in [0, k_{jam}]$ is the data grids for the numerical treatment. The speed-average density relationship before and after the HET is applied are given in Fig. 17 for each MPR value. Fig. 17 also gives the speed-density relationship for the critical RSS (bold red line) and background traffic flow (blue scatter dots). In Fig. 17-a, where the HET is not applied, as the MPR increases, the density also increase under the same speed. This is because

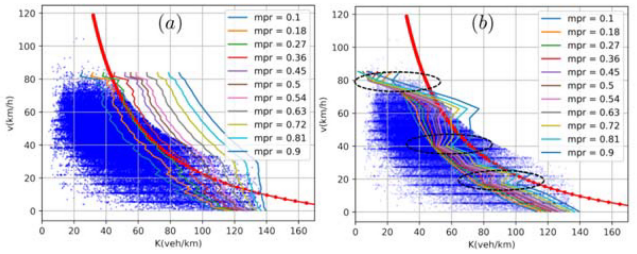


**FIGURE 17.** Speed-density relationship before (a) and after (b) the HET is applied.

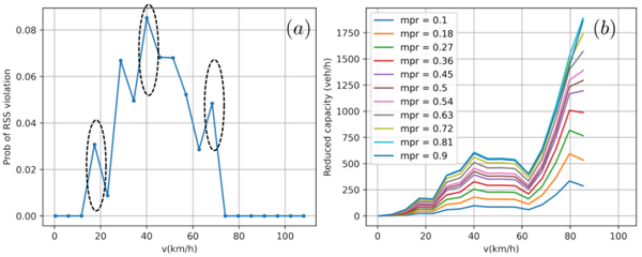


**FIGURE 18.** (a) Probability of the violation of RSS; (b) capacity reduction.

AVs can follow an AV or HDV with a short gap. In Fig. 17-b, the speed-average density relationship after the HET strategy is applied is given. We can see that there is considerable difference from the case in Fig. 17-a. For a given speed, the density when HET is applied drops. The decreasing magnitudes vary for different speed. This is because the HET only changes the density when the RSS violation occurs. When RSS violation doesn't occur, the density should not change (as demonstrated by speed-average density relationship near zero speed in Fig. 17-a and b). Three eclipse where the density decreases a lot are indicated, which coincide with the high RSS violation probability in Fig. 18. Fig. 18 is the ratio of the data points located at the right of the critical speed-density line specified by RSS (bold red line in Fig. 17). Due to the density decrease, the capacity is compromised. Fig. 18-b and Fig. 19 present the capacity loss and capacity loss ratio (compared with the case when HET is not applied, i.e., Fig. 17-a) due to the HET. At speed about 40km/h where the RSS violation probability is high, the capacity loss is about 600 veh/h, or 50%. Combining Fig. 18-b, Fig. 19 and Fig. 16, it is concluded that the proposed HET strategy sacrifice the capacity by reserving greater space gap, for the reduction of risk.

## VII. CONCLUSION AND REMARKS

Autonomous vehicle technology offers numerous advantages, including improved traffic safety, which is a crucial factor in the acceptance of AVs by consumers. However, in situations where AVs and human-driven vehicles coexist, traffic safety is affected by both AVs and HDVs. Therefore, it is necessary for the development of AV safety frameworks to consider the behavior of human drivers. The RSS model, which is a formal approach, establishes a minimum inter-vehicle distance for AV operations to prevent potential

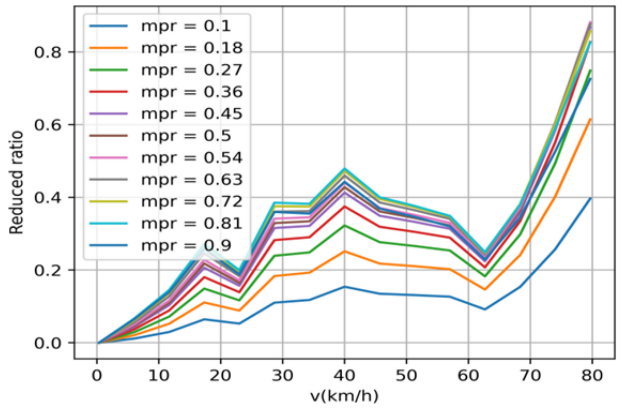


**FIGURE 19.** Reduced capacity ratio.

collisions. However, traditional RSS models do not take into account the risky behavior exhibited by human drivers, such as maintaining shorter and riskier distances. In scenarios where a leading vehicle suddenly brakes, an AV involved in the situation must choose either to enter the leader's crash zone or collide with a following vehicle, creating a dilemma. To address this, our research expands upon this scenario, introducing a generalized polylemma scenario, and proposes a HET strategy. By maintaining a greater inter-vehicle gap, the potential risks can be mitigated. Potential extensions of the research include the following aspects:

1) The current framework is based on deterministic assumptions. Real-world traffic flow operations are subject to significant stochasticity. As a result, further investigation is necessary to develop a robust RSS framework that can address both stochasticity and human error concurrently.

2) This study primarily focuses on the longitudinal dimension of vehicle movement. However, since vehicles operate in both longitudinal and lateral dimensions, a two-dimensional polylemma can be developed to accommodate a broader range of scenarios where human errors occur in a two-dimensional setting or in a coordinated platoon mode.

3) In mixed autonomous vehicle flow scenarios, human drivers may adapt their behavior, and it is essential to detect these changes in real-time to ensure traffic safety. Therefore, research can be conducted on determining the parameters of each human driver dynamically to support real-time traffic safety checks. Further, empirical investigation is necessary to validate the effectiveness of the proposed method.

## REFERENCES

[1] B. Peng, M. F. Keskin, B. Kulcsár, and H. Wymeersch, "Connected autonomous vehicles for improving mixed traffic efficiency in unsignalized intersections with deep reinforcement learning," *Commun. Transp. Res.*, vol. 1, Dec. 2021, Art. no. 100017.

[2] R. Yu, Y. Wang, Z. Zou, and L. Wang, "Convolutional neural networks with refined loss functions for the real-time crash risk analysis," *Transp. Res. Part C, Emerg. Technol.*, vol. 119, Oct. 2020, Art. no. 102740.

[3] T. Olovsson, T. Svensson, and J. Wu, "Future connected vehicles: Communications demands, privacy and cyber-security," *Commun. Transp. Res.*, vol. 2, Dec. 2022, Art. no. 100056, doi: 10.1016/j.commtr.2022.100056.

[4] J. Wu and X. Qu, "Intersection control with connected and automated vehicles: A review," *J. Intell. Connect. Veh.*, vol. 5, no. 3, pp. 260–269, 2022.

[5] M. Khayatian et al., "Cooperative driving of connected autonomous vehicles using responsibility-sensitive safety (RSS) rules," in *Proc. ACM/IEEE 12th Int. Conf. Cyber-Phys. Syst.*, 2021, pp. 11–20, doi: 10.1145/3450267.3450530.

[6] X. Qu, H. Lin, and Y. Liu, "Envisioning the future of transportation: Inspiration of ChatGPT and large models," *Commun. Transp. Res.*, vol. 3, Dec. 2023, Art. no. 100103, doi: 10.1016/j.commtr.2023.100103.

[7] X. Wang, D. Tang, S. Pei, P. Li, R. Yu, and K. Xie, "Transferability of urban arterial safety performance functions between Shanghai and Guangzhou, China," *J. Transp. Eng., Part A, Syst.*, vol. 148, no. 2, Feb. 2022, Art. no. 04021110, doi: 10.1061/JTEPBS.0000625.

[8] S. Sadraddini, S. Sivaranjani, V. Gupta, and C. Belta, "Provably safe cruise control of vehicular platoons," *IEEE Control Syst. Lett.*, vol. 1, no. 2, pp. 262–267, 2017.

[9] R. De Iaco, S. L. Smith, and K. Czarnecki, "Universally safe swerve maneuvers for autonomous driving," *IEEE Open J. Intell. Transp. Syst.*, vol. 2, pp. 482–494, 2021.

[10] R. Yu, Y. Zheng, and X. Qu, "Dynamic driving environment complexity quantification method and its verification," *Transp. Res. Part C, Emerg. Technol.*, vol. 127, Jun. 2021, Art. no. 103051.

[11] C. Chai, X. Zeng, X. Wu, and X. Wang, "Safety evaluation of responsibility-sensitive safety (RSS) on autonomous car-following maneuvers based on surrogate safety measurements," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, 2019, pp. 175–180. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8917421/

[12] Y. Peng, G. Tan, and H. Si, "RTA-IR: A runtime assurance framework for behavior planning based on imitation learning and responsibility-sensitive safety model," *Expert Syst. Appl.*, vol. 232, Dec. 2023, Art. no. 120824.

[13] B. Gassmann et al., "Towards standardization of AV safety: C++ library for responsibility sensitive safety," in *Proc. IEEE Intell. Veh. Symp. (IV)*, 2019, pp. 2265–2271. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8813885/

[14] P. Guo, Q. Zhu, and X. Wu, "Responsibility-Sensitive collision risk assessment and maneuvering safety evaluation," in *Proc. IEEE 25th Int. Conf. Intell. Transp. Syst. (ITSC)*, 2022, pp. 99–104. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9921955/

[15] X. Wang, C. Ye, M. Quddus, and A. Morris, "Pedestrian safety in an automated driving environment: Calibrating and evaluating the responsibility-sensitive safety model," *Accid. Anal. Prevent.*, vol. 192, Nov. 2023, Art. no. 107265.

[16] M. Hekmatnejad et al., "Encoding and monitoring responsibility sensitive safety rules for automated vehicles in signal temporal logic," in *Proc. 17th ACM-IEEE Int. Conf. Formal Methods Models Syst. Design*, 2019, pp. 1–11, doi: 10.1145/3359986.3361203.

[17] M. Naumann et al., "On responsibility sensitive safety in car-following situations-a parameter analysis on german highways," in *Proc. IEEE Intell. Veh. Symp. (IV)*, 2021, pp. 83–90. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9575420/

[18] P. Lin and M. Tsukada, "Cooperative path planning using responsibility-sensitive safety (RSS)-based potential field with sigmoid curve," in *Proc. IEEE 95th Veh. Technol. Conf.*, 2022, pp. 1–7. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9860508/

[19] D. Q. Tran and S.-H. Bae, "Improved responsibility-sensitive safety algorithm through a partially observable Markov decision process framework for automated driving behavior at non-signalized intersection," *Int. J Autom. Technol.*, vol. 22, no. 2, pp. 301–314, Apr. 2021, doi: 10.1007/s12239-021-0029-z.

[20] N. Lyu, J. Wen, Z. Duan, and C. Wu, "Vehicle trajectory prediction and cut-in collision warning model in a connected vehicle environment," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 2, pp. 966–981, Feb. 2022.

[21] S. Shalev-Shwartz, S. Shammah, and A. Shashua, "On a formal model of safe and scalable self-driving cars," 2018, *arXiv:1708.06374*.

[22] X. Qi, Y. Ni, Y. Xu, Y. Tian, J. Wang, and J. Sun, "Autonomous vehicles' car-following drivability evaluation based on driving behavior spectrum reference model," *Transp. Res. Rec.*, vol. 2675, no. 7, pp. 129–141, Jul. 2021, doi: 10.1177/0361198121994857.
[23] Z. Liu, H. Wu, and R. Li, "Effects of the penalty mechanism against traffic violations in China: A joint frailty model of recurrent violations and a terminal accident," *Accid. Anal. Prevent.*, vol. 141, Jun. 2020, Art. no. 105547.
[24] A. J. Khattak and B. Wali, "Analysis of volatility in driving regimes extracted from basic safety messages transmitted between connected vehicles," *Transp. Res. Part C, Emerg. Technol.*, vol. 84, pp. 48–73, Nov. 2017.
[25] Q. Yuan, X. Xu, M. Xu, J. Zhao, and Y. Li, "The role of striking and struck vehicles in side crashes between vehicles: Bayesian bivariate probit analysis in China," *Accid. Anal. Prevent.*, vol. 134, Jan. 2020, Art. no. 105324, doi: 10.1016/j.aap.2019.105324.
[26] K. Yang, C. A. Haddad, G. Yannis, and C. Antoniou, "Classification and evaluation of driving behavior safety levels: A driving simulation study," *IEEE Open J. Intell. Transp. Syst.*, vol. 3, pp. 111–125, 2022, doi: 10.1109/OJITS.2022.3149474.
[27] Z. Li, C. Xu, Y. Guo, P. Liu, and Z. Pu, "Reinforcement learning-based variable speed limits control to reduce crash risks near traffic oscillations on freeways," *IEEE Intell. Transp. Syst. Mag.*, vol. 13, no. 4, pp. 64–70, 2020.
[28] R. E. Stern et al., "Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments," *Transp. Res. Part C, Emerg. Technol.*, vol. 89, pp. 205–221, Apr. 2018.
[29] S. Singh, "Critical reasons for crashes investigated in the national motor vehicle crash causation survey," U.S. Dept. Transp., Nat. Highway Traffic Safety Admin., Washington, DC, USA, Rep. DOT HS 812 115, 2018. Accessed: Mar. 14, 2024. [Online]. Available: https://trid.trb.org/view.aspx?id=1346216&source=post_page
[30] O. Zheng, M. Abdel-Aty, Z. Wang, S. Ding, D. Wang, and Y. Huang, "AVOID: Autonomous vehicle operation incident dataset across the globe," 2023, arXiv.2303.12889.
[31] A. Abdelhalim and M. Abbas, "A real-time safety-based optimal velocity model," *IEEE Open J. Intell. Transp. Syst.*, vol. 3, pp. 165–175, 2022, doi: 10.1109/OJITS.2022.3147744.
[32] S. Kitajima, H. Chouchane, J. Antona-Makoshi, N. Uchida, and J. Tajima, "A nationwide impact assessment of automated driving systems on traffic safety using multiagent traffic simulations," *IEEE Open J. Intell. Transp. Syst.*, vol. 3, pp. 302–312, 2022, doi: 10.1109/OJITS.2022.3165769.
[33] L. A. Pipes, "An operational analysis of traffic dynamics," *J. Appl. Phys.*, vol. 24, no. 3, pp. 274–281, 1953.
[34] T. W. Forbes. "Human factor considerations in traffic flow theory: Highway research record." 1963. Accessed: Mar. 14, 2024. [Online]. Available: https://trid.trb.org/View/117024
[35] T. W. Forbes and M. E. Simpson, "Driver-and-vehicle response in freeway deceleration waves," *Transp. Sci.*, vol. 2, no. 1, pp. 77–104, Feb. 1968, doi: 10.1287/trsc.2.1.77.
[36] P. G. Gipps, "A behavioural car-following model for computer simulation," *Transp. Res. Part B, Methodol.*, vol. 15, no. 2, pp. 105–111, 1981.
[37] L. Elmorshedy, B. Abdulhai, and I. Kamel, "Quantitative evaluation of the impacts of the time headway of adaptive cruise control systems on congested urban freeways using different car following models and early control results," *IEEE Open J. Intell. Transp. Syst.*, vol. 3, pp. 288–301, 2022, doi: 10.1109/OJITS.2022.3166394.
[38] E. Andreotti, Selpi, and M. Aramrattana, "Cooperative merging strategy between connected autonomous vehicles in mixed traffic," *IEEE Open J. Intell. Transp. Syst.*, vol. 3, pp. 825–837, 2022, doi: 10.1109/OJITS.2022.3179125.
[39] M. Geisslinger, R. Trauth, G. Kaljavesi, and M. Lienkamp, "Maximum acceptable risk as criterion for decision-making in autonomous vehicle trajectory planning," *IEEE Open J. Intell. Transp. Syst.*, vol. 4, pp. 570–579, 2023, doi: 10.1109/OJITS.2023.3298973.
[40] G. Sidorenko, A. Fedorov, J. Thunberg, and A. Vinel, "Towards a complete safety framework for longitudinal driving," *IEEE Trans. Intell. Veh.*, vol. 7, no. 4, pp. 809–814, Dec. 2022.
[41] M. Luckcuck, M. Farrell, L. A. Dennis, C. Dixon, and M. Fisher, "Formal specification and verification of autonomous robotic systems: A survey," *ACM Comput. Surv.*, vol. 52, no. 5, pp. 1–41, Sep. 2020, doi: 10.1145/3342355.
[42] O. Hassanin, X. Wang, X. Wu, and X. Xu, "Efficiency performance and safety evaluation of the responsibility-sensitive safety in freeway car-following scenarios using automated longitudinal controls," *Accid. Anal. Prevent.*, vol. 177, Nov. 2022, Art. no. 106799.
[43] L. Li, X. Peng, F.-Y. Wang, D. Cao, and L. Li, "A situation-aware collision avoidance strategy for car-following," *IEEE/CAA J. Automatica Sinica*, vol. 5, no. 5, pp. 1012–1016, 2018.
[44] C. Chai, X. Zeng, X. Wu, and X. Wang, "Evaluation and optimization of responsibility-sensitive safety models on autonomous car-following maneuvers," *Transp. Res. Rec.*, vol. 2674, no. 11, pp. 662–673, Nov. 2020, doi: 10.1177/0361198120948507.
[45] G. Sidorenko, D. Plöger, J. Thunberg, and A. Vinel, "Emergency braking with ACC: How much does V2V communication help?" *IEEE Netw. Lett.*, vol. 4, no. 3, pp. 157–161, Sep. 2022.
[46] M. J. Kim and Y. M. Kim, "RSS model improvement considering road conditions for the application of a variable focus function camera," *Sensors*, vol. 23, no. 2, p. 592, 2023.
[47] S. Liu et al., "Calibration and evaluation of responsibility-sensitive safety (RSS) in automated vehicle performance during cut-in scenarios," *Transp. Res. Part C, Emerg. Technol.*, vol. 125, Apr. 2021, Art. no. 103037.
[48] X. Xu, X. Wang, X. Wu, O. Hassanin, and C. Chai, "Calibration and evaluation of the Responsibility-Sensitive Safety model of autonomous car-following maneuvers using naturalistic driving study data," *Transp. Res. Part C, Emerg. Technol.*, vol. 123, Feb. 2021, Art. no. 102988.
[49] H. Zheng et al., "Learning-based safe control for robot and autonomous vehicle using efficient safety certificate," *IEEE Open J. Intell. Transp. Syst.*, vol. 4, pp. 419–430, 2023, doi: 10.1109/OJITS.2023.3280573.
[50] K. Vellenga, H. J. Steinhauer, A. Karlsson, G. Falkman, A. Rhodin, and A. C. Koppisetty, "Driver Intention recognition: State-of-the-art review," *IEEE Open J. Intell. Transp. Syst.*, vol. 3, pp. 602–616, 2022, doi: 10.1109/OJITS.2022.3197296.
[51] D. I. Tselentis and E. Papadimitriou, "Driver profile and driving pattern recognition for road safety assessment: Main challenges and future directions," *IEEE Open J. Intell. Transp. Syst.*, vol. 4, pp. 83–100, 2023, doi: 10.1109/OJITS.2023.3237177.
[52] R. Trauth, K. Moller, and J. Betz, "Toward safer autonomous vehicles: occlusion-aware trajectory planning to minimize risky behavior," *IEEE Open J. Intell. Transp. Syst.*, vol. 4, pp. 929–942, 2023, doi: 10.1109/OJITS.2023.3336464.
[53] L. Vasile, N. Dinkha, B. Seitz, C. Däsch, and D. Schramm, "Comfort and safety in conditional automated driving in dependence on personal driving behavior," *IEEE Open J. Intell. Transp. Syst.*, vol. 4, pp. 772–784, 2023, doi: 10.1109/OJITS.2023.3323431.
[54] A. Scamarcio, P. Gruber, S. De Pinto, and A. Sorniotti, "Anti-jerk controllers for automotive applications: A review," *Annu. Rev. Control*, vol. 50, pp. 174–189, Jul. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1367578820300237
[55] M. Makridis, K. Mattas, and B. Ciuffo, "Response time and time headway of an adaptive cruise control. An empirical characterization and potential impacts on road capacity," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 4, pp. 1677–1686, Apr. 2020.