Submitted to *Transportation Science*
manuscript (Please, provide the mansucript number!)

A microscopic investigation into the capacity drop: impacts of longitudinal behavior on the queue discharge rate

Kai Yuan, Victor L. Knoop, Serge P. Hoogendoorn
Faculty of Civil Engineering and Geoscience, Delft University of Technology, 2628 CN Delft, The Netherlands, K.Yuan@tudelft.nl

The capacity drop indicates that the queue discharge rate is lower than the free-flow capacity. Studies show that the queue discharge rate varies under different traffic conditions. Empirical data show that the queue discharge rate increases as the speed in congestion increases. Insights into the underlying behavioral mechanisms that cause these variable queue discharge rates can help minimize traffic delays and eliminate congestion. However, to the best of the authors knowledge, few efforts have been devoted to testing impacts of traffic behaviors on the queue discharge rate. This paper tries to fill this gap. We investigate to what extent the acceleration spread and reaction time can influence the queue discharge rate. It is found that the (inter-driver) acceleration spread does not reduce the queue discharge rates as much as found empirically. Modelling reaction time might be more important than modeling acceleration for capacity drop in car-following models. A speed-dependent reaction time mechanism for giving variable queue discharge rates is proposed. That is, decreasing reaction time as the speed in congestion increases can give the same queue discharge rate as found empirically. This research suggests that motivating drivers to speed up earlier could increase the queue discharge rate and thereby minimize delays.

Key words: capacity drop; bounded acceleration; reaction time; analytical modeling
History:

1. Introduction.

Road congestion can be categorized into two classes: standing queues with heads fixed at a bottleneck and stop-and-go waves with queue fronts moving upstream. The bottleneck is a fixed point where the congestion head is located. Once congestion sets in, the flow out of congestion is the queue discharge rate. This flow is generally lower than the free-flow capacity, i.e., the maximum flow which is also referred to pre-queue capacity in literature. This phenomenon is called the capacity drop. For simplicity, we use capacity for short to stand for the free-flow capacity in this manuscript.

The magnitude of the capacity drop is not constant. Empirical data show that the queue discharge rate varies considerably at the same location (Chung et al. 2007, Yuan et al. 2016). This is shown to correlate well with congestion states (Yuan et al. 2015, Oh and Yeo 2015). Yuan et al. (2015) reveal a linear relation between the speed in congestion and the queue discharge rate (see Figure 1). The specific relation is based on empirical data collected on freeway A4 and A12 in the Netherlands. Road design and control measures can contribute to varying queue discharge rates (Srivastava and Geroliminis 2013, Oh and Yeo 2015). These findings show that there might be promising strategies that can increase the queue discharge rate to reduce delays. However, to determine effective approaches, an insight is needed into the underlying behavioral mechanisms that cause the capacity drop. Therefore, this paper tries to investigate the impacts of driver behavior on the queue discharge rate.

The capacity drop is usually observed at fixed bottlenecks, especially merging bottlenecks (e.g., on-ramp and lane-drop bottlenecks). Before the standing queue forms (i.e., before a break down) at a fixed merging bottleneck, loop detectors can measure flow approaching the capacity. So empirical data at fixed merging bottlenecks can show the reduction from the capacity to the queue discharge rate, i.e., the capacity drop phenomenon. However, a reduction of capacity does not only occur at bottlenecks. In fact, some research (Kerner 1998, Yuan et al. 2016, Oh and Yeo 2015) report on queue discharge rates out of stop-and-go waves on homogeneous freeway sections, which are also lower than the capacity. In this paper, we also call that a capacity drop. Indeed, this drop cannot be detected directly in this spatial-temporal traffic flow situation. When understanding the capacity drop from a driver behavioral perspective, this manuscript categorizes the mechanism of the capacity drop into lateral mechanisms (i.e., the capacity drop due to lane changing behavior) and longitudinal mechanisms (i.e., the capacity drop due to longitudinal behaviors).

By empirically analyzing NGSIM trajectory datasets, Oh and Yeo (2015) firstly shows that the contribution of traffic passing stop-and-go waves to discharge rates outperforms the contribution of lane change related traffic to discharge rates. Secondly, it is found that the queue discharge rate could be related to the severity of congestion in the absence of lane changing. Our work can hence be seen in the perspective shaped by Oh and Yeo (2015), and we therefore focus on the study on the drivers longitudinal behavior mechanism behind the capacity drop.

As shown in the literature review, there are two categories of longitudinal behaviors which can reduce the queue discharge rate, that is, intra-driver (variable reaction time) and inter-driver (acceleration spread) variable behaviors. This contribution investigate to what extent these two behaviors contribute to the capacity drop. The acceleration can give the capacity drop with inter-driver acceleration spread. Inter-driver acceleration spread (or in short: acceleration spread) means

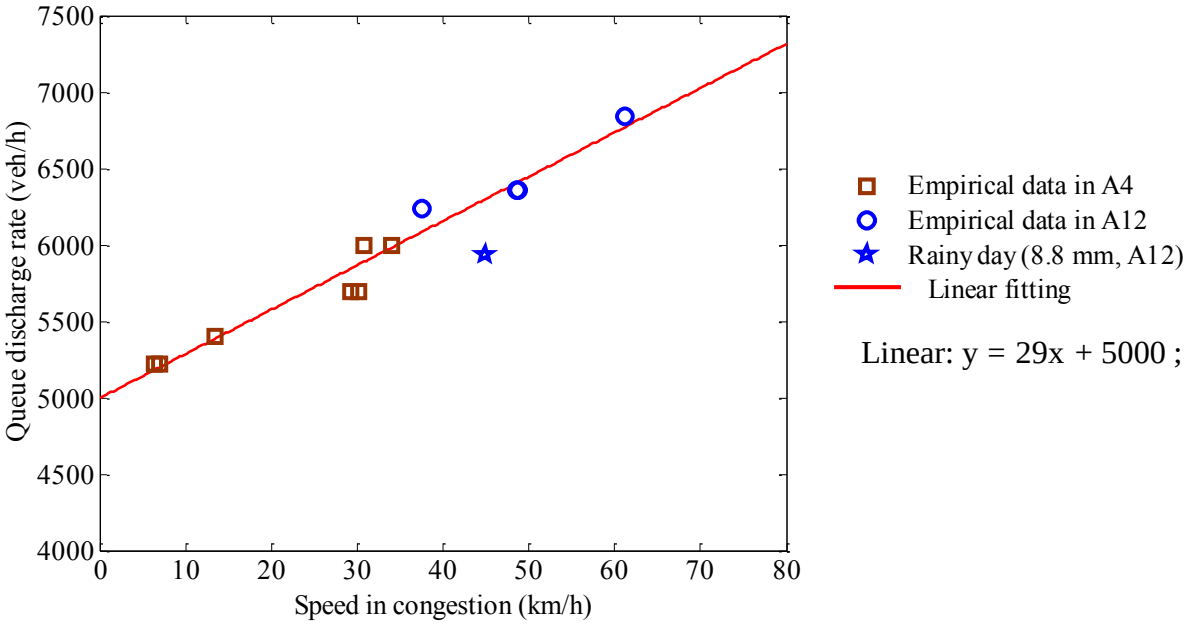


Figure 1 **Empirical data showing a relation between the speed in congestion and the queue discharge rate (Yuan et al. 2015).**

that vehicles do not have the same acceleration. As a result, voids will be created between a low-acceleration vehicle and its high-acceleration predecessor. The reaction time indicates how long a following vehicle needs to take to react to the change of its leaders driving behavior. Voids can also be created if the followers reaction time is longer than Newells reaction time (see section 3.2). In this paper, we call such long reaction time the extended reaction time. To what extent the inter-driver acceleration spread and the extended reaction time contribute to the capacity drop is unknown. In this paper, we study the impacts of the acceleration spread and the extended reaction time on the queue discharge rate.

This paper develops analytical models to investigate the independent impact of accelerations and reaction time. Furthermore, we design numerical experiments for two objectives. First, the experiment is used to validate the analytical model to ensure the approximation in the model is accurate enough. Second, we use the experiment to see the combination effects of acceleration spread and reaction time on the queue discharge rate. The empirical relation revealed in Yuan et al. (2015) is the reference used in our analyses, see Figure 1.

Our study excludes several factors that may influence the queue discharge rate. Firstly, drivers perspectives, i.e., whether drivers are aggressive or timid, are excluded. Secondly, lane changing is not considered in this paper. As argued in Papageorgiou et al. (2008), if we simulate a stop-and-go wave moving on a homogeneous road section, lane changing frequency should be very low in an acceleration mode.

The outline of the paper is as follows: we start with a literature review in section 2. Then section 3 presents the analytical investigation on the capacity drop. In section 4, we use simulations to validate the analytical model (section 4.2) and investigate the combination of acceleration and reaction time (section 4.3), followed by discussions and conclusions in section 5.

2. Literature Review.

A wide range of capacity drop values have been observed, which are reviewed in section 2.1. The wide range of the capacity drop values could be due to various queue discharge rates which correlate well with different congested states. The research objective of this paper is to investigate the relation between driving behavior and the queue discharge rate. Hence, section 2.2 reviews previous traffic behavioral mechanism of the capacity drop.

2.1. Empirical features of the capacity drop.

The capacity drop was reported in 1991 with a drop of 6% (Hall and Agyemang-Duah 1991) and 3% (Banks 1991). In the past decades, the capacity drop have been studied more often, with values of the drop ranging between 3% and 18% (Oh and Yeo 2015). In Cassidy and Bertini (1999) and Bertini and Leal (2005) the capacity drop ranges from 8% to 10%. In Srivastava and Geroliminis (2013) the capacity falls by 15% at an on-ramp bottleneck. Chung et al. (2007) show a range of capacity drop from 3% to 18% with data collected at three active bottlenecks, which shows a drop from 8% to 18% at the same location. Cassidy and Rudjanakanoknad (2005) observe capacity drop between 8.3% and 14.7%.

We argue that the wide range of capacity drop values in literature correlates well with the congestion state. Yuan et al. (2015) show a positive correlation between the queue discharge rate and the speed in congestion with empirical data collected on freeways in the Netherlands. Oh and Yeo (2015) find that the queue discharge rate is related to the severity of congestion by analyzing microscopic trajectory data. Hence, the research question is: what is the mechanism behind the dependency of discharge rate on the congested states? Answering this question might help to better understand the microscopic mechanism of the capacity drop.

2.2. Overview of Assumptions on Mechanisms of the Capacity Drop.

Many studies have been reporting the capacity drop in the past decades. Table 1 summarizes most of the existing most popular assumptions on the traffic behavioral mechanism of capacity drop. Generally, we can divide them into three categories: bounded acceleration capability, inter-driver/vehicle spread, and intra-driver spread.

Bounded acceleration capability means vehicles cannot accelerate instantaneously. Consequently lane change maneuvers can create voids in the traffic stream (Laval and Daganzo 2006, Duret et al.

Table 1 Simulation result

Basis mechanisms	Assumptions on mechanisms	References
(a) Bounded acceleration capability	Lane changing	Laval and Daganzo (2006) Duret et al. (2010) Leclercq et al. (2011) Leclercq et al. (2015) Coifman and Kim (2011)
(b) Inter-driver/vehicle spread	Acceleration spread	Papageorgiou et al. (2008)
	Multi-class vehicles	Wong and Wong (2002)
(c) Intra-driver spread	Variance-driven time headways	Treiber et al. (2006)
	Multiphase car-following theory	Zhang and Kim (2005)
	Asymmetric driving behavior theory	Yeo (2008)
	Activation level	Tampère (2004)
	Drivers' perspectives	Chen et al. (2014)

2010, Leclercq et al. 2011, 2015). The limited acceleration causes that the lane changing vehicle cannot catch up with its new predecessor. Coifman and Kim (2011) show that lane changing in the far downstream of the congestion can result in the capacity drop, too. Insertions result in shock waves in the new lane and the divergences in the old lane create voids which cannot be filled in duo to the bounded acceleration capability. So an aggregated flow detected in the downstream of queue could be lower than the capacity.

Inter-driver/vehicle spread indicates the spread of drivers and vehicles. Papageorgiou et al. (2008) state that the capacity drop is due to the acceleration difference between two successive vehicles. Voids can be created between a low-acceleration vehicle and its high-acceleration predecessor. Wong and Wong (2002) reproduce the capacity drop in numerical simulations when formulating a multi-class traffic flow model as an extension of the Lighthill, Whitham and Richards (LWR) model (Lighthill and Whitham 1955, Richards 1956) with heterogeneous drivers.

The third popular explanation, intra-driver spread, assumes driver behavior varies depending on traffic conditions. Treiber et al. (2006) assume drivers would choose a longer time headway in congestion than that in free flow. The preferred time headway in congestion increases as density increases. This assumption, also called variance-driven time headways, is based on an empirical observation of an increasing time gaps between one vehicle's front bumper and the rear bumper of the preceding vehicle after a considerable queuing time in Nishinari et al. (2003). Zhang and Kim

(2005) propose a multi-phase car-following traffic flow theory to reproduce the capacity drop. They highlight that the capacity drop is a result of driver behavior spread across three phases, i.e., acceleration, deceleration and coasting. Yeo (2008) validates the acceleration and deceleration curves to further develop the asymmetric microscopic traffic flow theory based on empirical trajectory data, explaining the capacity drop as a difference of the maximum flow between the acceleration and the deceleration curve in density-flow fundamental diagram. The asymmetric driver behavior theory is also applied in Oh and Yeo (2015) to understand the impacts of stop-and-go waves on the capacity drop. Tampère (2004) assumes drivers behavior depends on a temporary, traffic condition dependent variable activation level. The low activation level used to accounted for a loss of motivation. They reproduce the capacity drop as a result of low activation level in case studies. Chen et al. (2014) explain the capacity drop as a result of the change of aggressiveness, that is, the intra-driver behavior spread is described as four different reaction (i.e., concave, convex, non-decreasing and constant) patterns to disturbances. Those different reaction patterns, which are validated in NGSIM dataset in (Chen et al. 2012), can give capacity drop ranging from 8% to 23% in simulations.

In this paper, we focus on studying the impacts of acceleration spread and the reaction time extension on the queue discharge rate reduction and its correlation with the congested states.

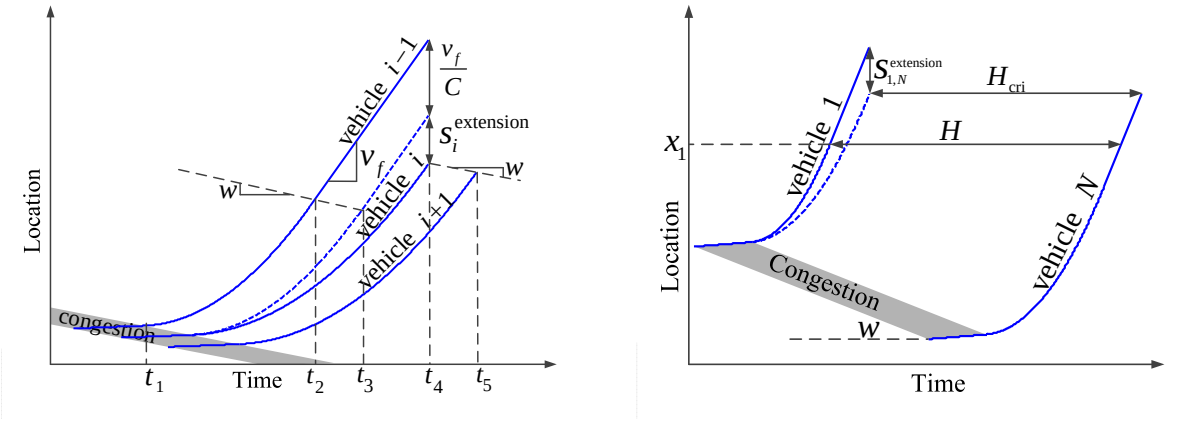
3. Analytical Investigation.

This section analytically investigates to what extent the acceleration spread (Section 3.1) and reaction time extension (Section 3.2) can independently account for the capacity drop. In each of Section 3.1 and Section 3.2, we firstly present a numerical expression of the queue discharge rate, followed by analysis of the model properties.

Using mathematical derivations shows that the capacity drop as consequence of acceleration spread is smaller than that found in empirical studies. An intra-driver reaction time extension mechanism can model similar queue discharge rates as reality. For practical purpose, these conclusions indicate that pushing slowly driving vehicles to speed up earlier, rather than managing vehicular acceleration, might be an approach for minimizing capacity drops and delays.

3.1. Capacity drop due to acceleration spread.

In this section we derivate analytical formulas for the capacity drop in Section 3.1.1 and find the acceleration spread does not give sufficient queue discharge rate reduction compared to empirical observations in Section 3.1.2.



(a) Spacing extensions due to acceleration variability.

(b) Queue discharge rates measurements with acceleration spread.

Figure 2 Measurements of the capacity drop due to the acceleration spread.

3.1.1. Analytical expressions of queue discharge rates Let us consider a stop-and-go wave moving upstream on a homogeneous road section, see Figure 2. The grey block is the stop-and-go wave and bold lines are vehicular trajectories. The traffic in the scenario is described by a triangular fundamental diagram with a wave speed $-w$, free-flow speed v_f and capacity C. The critical density and maximum jam density are given by ρ_{cri} and ρ_{jam}, respectively. There are N vehicles in total in the queue in a single lane, obeying the first-in-first-out (FIFO) rule. Each vehicle is numbered $i(i=1,2,...N)$, increasing from the head of the queue $(i=1)$ to the tail $(i=N)$. The speed and density in the queue are v_j and ρ_j , respectively. When all vehicles reach the free-flow speed after leaving the queue, the free-flow spacing and time headway between vehicle i and $i-1$ is given by s_i and h_i, respectively. The minimum free-flow spacing s_{min} for all vehicles should be $\frac{1}{\rho_{\text{cri}}}$ (or the minimum time headway $h_{\text{min}}=\frac{1}{C}$), indicating no capacity drop at all. Each vehicle i is described by two constants, its desired acceleration a_i^{desire} and acceleration a_i. In principle, every vehicle accelerates with its desired acceleration. However, s_i is at the low end bounded by s_{min}. Therefore, if $a_i=a_i^{\text{desire}}$ will result in $s_i<s_{\text{min}}$, we set $a_i<a_i^{\text{desire}}$ to ensure $s_i=s_{\text{min}}$. Note that $a_1=a_1^{\text{desire}}$. Desired accelerations fall within the interval $[a_{\text{min}},a_{\text{max}}]$. The reaction time of vehicle i is denoted as t_r. All vehicles have reached free-flow speed at x_1 where the sum of free-flow time headways from the second vehicle to the last vehicle is denoted as H, see Figure 2(b). We use H_{cri} to denote the sum of free-flow time headways from vehicle 2 to vehicle N in no-capacity-drop condition, i.e., $H_{cri}=\frac{n-1}{C}$.

If all vehicles follow the continuous Newell car-following model (Newell 2002), constructed by shifting its predecessors trajectory by spacing $\Delta s=\frac{1}{\rho_{\text{jam}}}$ and time $\Delta t=\frac{\Delta s}{w}=\frac{1}{w\rho_{\text{jam}}}$, there is no capacity drop.

It is impossible that all vehicles have the same acceleration. We assume the desired acceleration follows an uniform distribution bounded by $a_{\min}$ and $a_{\max}$, i.e., $a_{\text{desire}} \sim U[a_{\min}, a_{\max}]$. The uniform distribution is denoted by U. We exclude the impact of reaction time extensions by setting $t_r = \Delta t$. When the desired acceleration of vehicle i is higher than its leaders acceleration, we set $a_i = a_{i-1}$ to ensure the follower can neither overtake nor be too close ($s_i < s_{\min}$) to its leader. Otherwise, $a_i = a_i^{\text{desire}}$.

A void is created between two successive vehicles if the followers desired acceleration is lower than the predecessors acceleration. In Figure 2(a), a dashed line is the Newell trajectory of vehicle i. Note the void between the Newell trajectory and the real trajectory of vehicle i. The void indicates that the free-flow spacing is extended by $s_i^{\text{extension}}$:

$$s_i^{\text{extension}} = \frac{(v_j - v_f)^2 (a_{i-1} - a_i)}{2a_{i-1}a_i} = \frac{1}{2}(v_j - v_f)^2 \left(\frac{1}{a_i} - \frac{1}{a_{i-1}}\right) \tag{1}$$

Now let us consider N vehicles within a stop-and-go wave as in Figure 2(b). The average queue discharge rate $E(q_d)$ is expressed as:

$$E(q_d) = \frac{n-1}{E(H)} \tag{2}$$

$E(H)$ is the mean value of H. Now let us calculate $E(H)$ in order to calculate $E(q_d)$.

In Figure 2(b), the dashed line is a shifted trajectory from the trajectory of vehicle N by spacing $(n-1)\Delta s$ and time $(n-1)\Delta t$. According to Equation (1), the void between the trajectory of vehicle 1 and the dashed trajectory is $s_{1,N}^{\text{extension}} = \frac{1}{2}(v_j - v_f)^2 \left(\frac{1}{a_N} - \frac{1}{a_1}\right)$. At location x_1, all vehicles have reached the free-flow speed. So vehicle N reaches x_1 after the dashed trajectory has passed x_1 for $H_{cri} = \frac{n-1}{C}$. From Figure 2(b), we derive

$$E(H) = E\left(H_{cri} + \frac{s_{1,N}^{\text{extension}}}{v_f}\right) = \frac{n-1}{C} + \frac{1}{2v_f}(v_j - v_f)^2 \left[E\left(\frac{1}{a_N}\right) - E\left(\frac{1}{a_1}\right)\right] \tag{3}$$

which indicates that the calculation of $E\left(\frac{1}{a_N}\right)$ and $E\left(\frac{1}{a_1}\right)$ can give $E(q_d)$. For $E\left(\frac{1}{a_1}\right)$, since $a_1 = a_1^{\text{desire}} \sim U[a_{\min}, a_{\max}]$, we can have:

$$E\left(\frac{1}{a_1}\right) = \frac{\ln\left(\frac{a_{\max}}{a_{\min}}\right)}{a_{\max} - a_{\min}} \tag{4}$$

Now we need $E\left(\frac{1}{a_N}\right)$. Let $\left(a_{(1)}^{\text{desire}}, ..., a_{(N)}^{\text{desire}}\right)$ denote the corresponding order statistics of the random sample $(a_1^{\text{desire}}, ..., a_N^{\text{desire}})$ so that $a_{(1)}^{\text{desire}} \leq a_{(2)}^{\text{desire}} \leq ... \leq a_{(N)}^{\text{desire}}$. Because $a_N = a_{(1)}^{\text{desire}}$, the probability density function of a_N equals to the probability density function of the smallest order statistic $a_{(1)}^{\text{desire}}$. Thanks to order statics (Reiss 1989), we have the probability distribution function f_A of $a_{(1)}^{\text{desire}}$:

$$f_A\left(a_{(1)}^{\text{desire}}\right) = N\left(1 - F\left(a_{(1)}^{\text{desire}}\right)\right)^{N-1} f\left(a_{(1)}^{\text{desire}}\right) \tag{5}$$

where $F\left(a_{(1)}^{\text{desire}}\right)$ is the cumulative distribution function of the smallest desired acceleration and $f\left(a_{(1)}^{\text{desire}}\right)$ is the probability distribution function of the smallest desired acceleration. For $a_{(1)}^{\text{desire}} \sim U[a_{\min}, a_{\max}]$, we get $F\left(a_{(1)}^{\text{desire}}\right) = \frac{a^{\text{desire}} - a_{\min}}{a_{\max} - a_{\min}}$ and $f\left(a_{(1)}^{\text{desire}}\right) = \frac{1}{a_{\max} - a_{\min}}$. So the probability distribution function f_N of a_N can be given from (5) to:

$$f_N(a_N) = f_A\left(a_{(1)}^{\text{desire}}\right) = \left(\frac{a_{\max} - a_N}{a_{\max} - a_{\min}}\right)^{N-1} \frac{N}{a_{\max} - a_{\min}}, \text{ for } \mathrm{a}_{(1)}^{\text{desire}} = \mathrm{a_N} \in [\mathrm{a_{min}}, \mathrm{a_{max}}] \tag{6}$$

Setting $g(x) = \frac{1}{x}$, we estimate the second-order approximation of $E(g(a_N))$ with the Delta method:

$$E\left(\frac{1}{a_N}\right) = E(g(a_N)) \approx g(E(a_N)) + \frac{1}{2} g''(E(a_N))\, \sigma^2(E(a_N)) \tag{7}$$

where $E(a_N)$ and $\sigma^2(E(a_N))$ are the mathematical expectation and the standard deviation of a_N, which can be deduced from Equation (6). That is, $E(a_N) = \frac{a_{\max} + N a_{\min}}{1+N}$ and $\sigma^2(a_N) = \frac{N(a_{\max} - a_{\min})^2}{(2+N)} + \frac{a_{\max}^2(-N+1) + 2N a_{\min} a_{\max}}{(1+N)} - \frac{(a_{\max} + N a_{\min})^2}{(1+N)^2}$. Since $g''(E(a_N)) = 2\left(\frac{1+N}{a_{\max} + a_{\min} \cdot N}\right)^3$, the estimation of $E\left(\frac{1}{a_N}\right)$ is:

$$\begin{aligned} E\left(\frac{1}{a_N})\right) \approx & \frac{1+N}{a_{\max} + N a_{\min}} + \\ & \left(\frac{1+N}{a_{\max} + N a_{\min}}\right)^3 \left(\frac{N(a_{\max} - a_{\min})^2}{2+N} + \frac{a_{\max}^2(-N+1) + 2N a_{\max} a_{\min}}{1+N} - \left(\frac{a_{\max} + N a_{\min}}{1+N}\right)^2\right) \end{aligned} \tag{8}$$

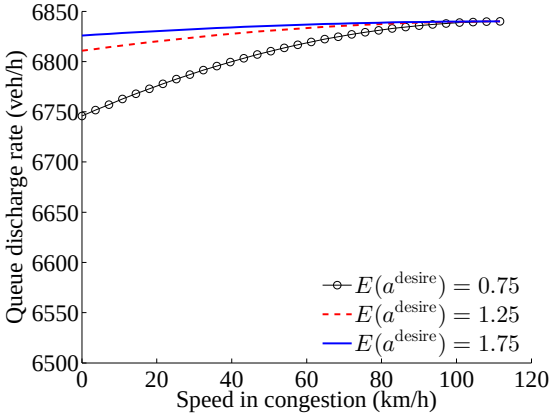
Incoporating Equation (4) and Equation (8) into Equation (3), we get $E(H)$:

$$\begin{aligned} E(H) = & \frac{N-1}{C} + \frac{(v_j - v_f)^2 (1+N)}{2 v_f (a_{\max} + a_{\min} \cdot N)} - \frac{(v_j - v_f)^2 \ln\left(\frac{a_{\max}}{a_{\min}}\right)}{2 v_f (a_{\max} - a_{\min})} + \\ & \frac{(1+N)^3 (v_j - v_f)^2}{2 v_f (a_{\max} + a_{\min} N)^3} \left(\frac{N(a_{\max} - a_{\min})^2}{2+N} + \frac{a_{\max}^2(-N+1) + 2N a_{\max} a_{\min}}{1+N} - \left(\frac{a_{\max} + N a_{\min}}{1+N}\right)^2\right) \end{aligned} \tag{9}$$

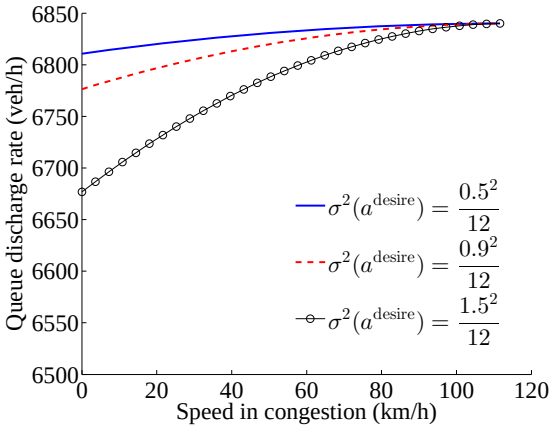
Incorporating Equation (9) into Equation (2) can give the estimated average queue discharge rate $E(q_d)$.

3.1.2. Analysis of model properties We set a triangular fundamental diagram with $w =$ 18km/h, $v_f = 114$km/h, $C = 6840$veh/h, $\rho = 60$veh/km and $\rho_{\text{jam}} = 440$km/h. This fundamental diagram indicates a similar traffic situation as that in Yuan et al. (2015). Different bounds for accelerations have been reported: for instance 0.5m/s^2 - 3m/s^2 (Leclercq et al. 2011), or 1.5m/s^2 - 2m/s^2 (Lebacque 2003). We combine these and set the limits for desired accelerations from 0.5m/s^2 to 2m/s^2. We will limit the range further.

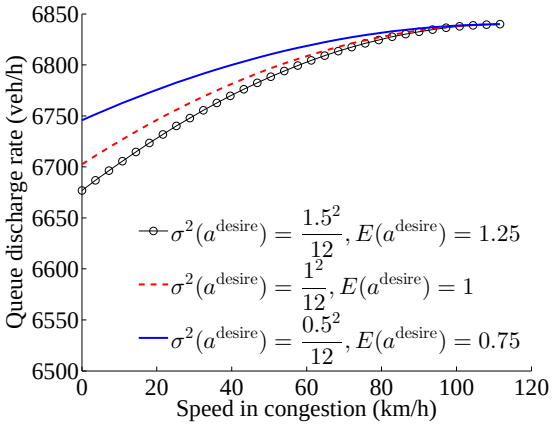
Consider a stop-and-go wave that propagates at speed w for $\tau = 10$ minutes. Variational theory (Daganzo 2005) gives the number of vehicles in the queue $N = \left\lfloor \frac{w \rho_{\text{jam}} \tau}{60} \right\rfloor = 1320$veh. This section analyses the queue discharge rate for this queue.



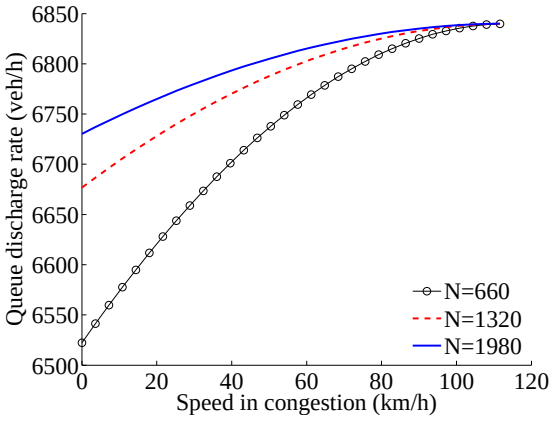
(a) Sensitivity of the analytical model to the mathematic expectation of desired accelerations.



(b) Sensitivity of the analytical model to the standard deviations of desired accelerations.



(c) Comparisons of impacts on queue discharge rates between the mean and deviation of desired accelerations. $a_{\min}$ is fixed to be 0.5 m/s^2 and $a_{\max}$ decreases from 2 m/s^2 to 1 m/s^2.



(d) Sensitivity of the analytical model to vehicle numbers.

Figure 3 Sensitivity of queue discharge rates when capacity drop is due to the acceleration spread.

As shown in Equation (9), $E(H)$ is a function of $a_{\min}$, $a_{\max}$ and N. The sensitivity of the queue discharge rate to the average desired accelerations, standard spread of desired accelerations and number of vehicles are evaluated with Equation (2) and (9), presented in Figure 3.

Figure 3(a) presents a relation between the speed in congestion v_j and the queue discharge rate q_d when setting $E(a^{\text{desire}})$ as 0.75m/s^2, 1.25m/s^2, 1.75m/s^2 respectively and $\sigma^2(a^{\text{desire}}) = \frac{0.5^2}{12}$m/s^2. We obtain so by setting the pair $(a_{\min}, a_{\max})$ to (0.5m/s^2, 1m/s^2), (1m/s^2, 1.5m/s^2) and (1.5m/s^2, 2m/s^2). We see that the faster the average desired acceleration, the higher the queue discharge rate.

Figure 3(b) presents the relation between v_j and q_d when $\sigma^2(a^{\text{desire}})$ equals to $\frac{0.5^2}{12}$m/s^2, $\frac{0.9^2}{12}$m/s^2 and $\frac{1.5^2}{12}$m/s^2, setting $E(a^{\text{desire}}) = 1.25$m/s^2. That is, the pair $(a_{\min}, a_{\max})$ are chosen to be

(1m/s^2,1.5m/s^2), (0.8m/s^2,1.7m/s^2) and (0.5m/s^2,2m/s^2) respectively. It indicates that the larger the spread, the lower the queue discharge rate.

If we fix $a_{\min} = 0.5\text{m/s}^2$ and decrease $a_{\max}$ from 2m/s^2 to 1m/s^2, then both of $E\left(a^{\text{desire}}\right)$ and $\sigma^2\left(a^{\text{desire}}\right)$ decreases. Figure 3(c) shows that the decrease of $a_{\max}$ increases the queue discharge rates. Since the decrease of $E\left(a^{\text{desire}}\right)$ and $\sigma^2\left(a^{\text{desire}}\right)$ will decrease and increase the queue discharge rate respectively, the increase of queue discharge rates in Figure 3(c) indicates that $\sigma^2\left(a^{\text{desire}}\right)$ has more influences on the queue discharge rate than $E\left(a^{\text{desire}}\right)$.

Figure 3(d) shows the sensitivity to N with $a_{\min} = 0.5\text{m/s}^2$ and $a_{\max} = 2\text{m/s}^2$. The more vehicles, the higher queue discharge rates. It is not a surprise because the followers acceleration is always limited by its leaders acceleration, that makes the acceleration spread decrease as the vehicle number increases. Since $N = 1320$veh means the congestion only propagates for 10min, the queue discharge rate can be even higher when setting a longer time of congestion propagation.

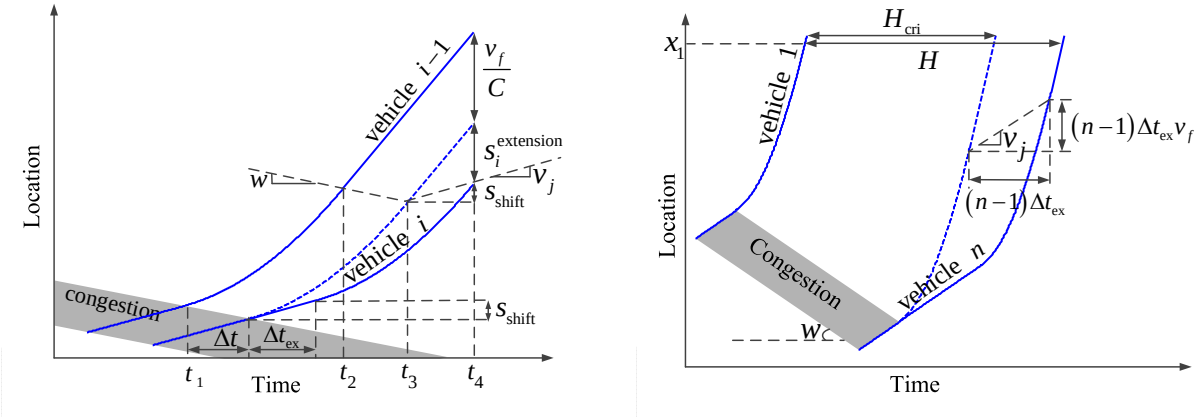
Setting $a_{\min} = 0.5\text{m/s}^2$, $a_{\max} = 2\text{m/s}^2$ and $N = 660$veh gives a considerable influence of the acceleration spread on queue discharge rates, shown as the line with circles in Figure 3(d). However, the contribution of acceleration spread to the queue discharge rate reduction is still marginal. In Figure 3(d) when $v_j = 0\text{km/h}$, the minimum queue discharge rate (6522veh/h) is still much higher than the empirical value (5000veh/h) shown in Figure 1.

Note that the hypothesis about the uniform desired acceleration distribution has already maximized the $\sigma^2\left(a^{\text{desire}}\right)$. In reality, the desired acceleration could follow some distribution with peaks (such as shown in Koutsopoulos and Farah (2012)) which will have smaller $\sigma^2\left(a^{\text{desire}}\right)$. Therefore, we can conclude that the acceleration spread is not a dominant factor for the capacity drop.

3.2. Capacity drop due to reaction time extension.

This section shows that the reaction time extension can considerably influence the queue discharge rate. A negative relation between the reaction time and the speed in congestion could result in a similar queue discharge rates as empirical findings. We give analytical expressions of queue discharge rates and the sensitivity analyses in section 3.2.1 and 3.2.2, respectively.

3.2.1. Analytical expressions of queue discharge rates A reaction time longer than Δt will lead to a larger spacing than $\frac{1}{\rho_{\text{cri}}}$, so the queue discharge rate will be lower than the capacity which is the capacity drop. Therefore, we consider only the cases when the reaction time is longer, and we define $t_r = \Delta t + \Delta t_{\text{ex}}$ where Δt is considered as a fixed reaction time (related to the fundamental diagram) and Δt_{ex} as a reaction time extention. As shown in Figure 4(a), two bold solid lines are trajectories of two successive vehicles accelerating from speed v_j up to free speed v_f. The followers reaction time is extended by Δt_{ex} from Δt. The dashed line is the followers trajectory when $\Delta t_{\text{ex}} = 0\text{s}$. The follower's trajectory can be considered as a shifted trajectory



(a) Spacing extensions due to reaction time extensions.

(b) Queue discharge rates measurements with reaction time extensions.

Figure 4 Measurements of the capacity drop due to the reaction time extension.

from the dashed line in time (by Δt_{ex}) and space (by s_{shift}). Geometry (see Figure 4(a)) gives us $s_i^{\text{extension}} = (v_f - v_j)\,\Delta t_{\text{ex}}$. So the free-flow spacing is extended to:

$$s_i = \frac{1}{\rho_{\text{cri}}} + s_i^{\text{extension}} = \frac{v_f}{C} + (v_f - v_j)\,\Delta t_{\text{ex}} \tag{10}$$

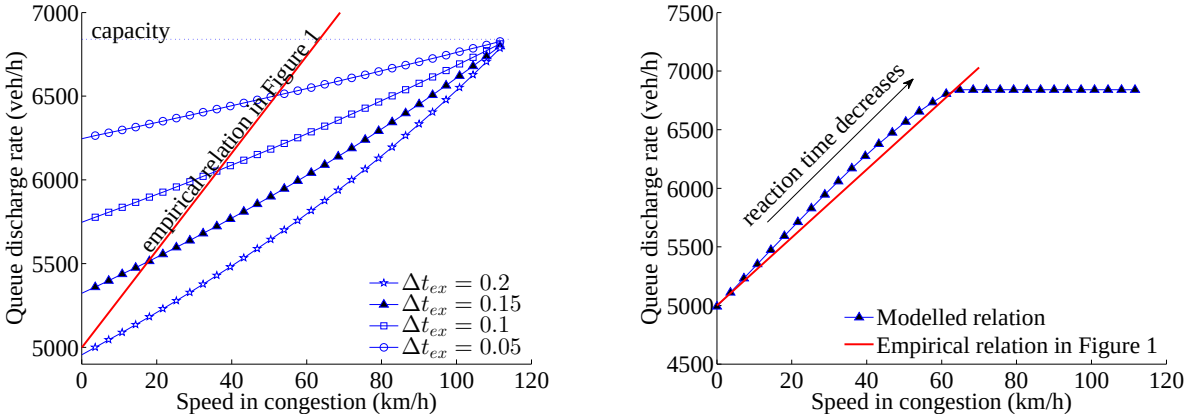
Considering N vehicles accelerating from a queue with the same acceleration (see Figure 4(b)), we can calculate the spacing between the first and the last vehicle as follows:

$$s_{1,N} = H_{\text{cri}} v_f + \sum_{i=1}^{N-1} s_i^{\text{extension}} = \frac{N-1}{\rho_{\text{cri}}} + (N-1)\,(v_f - v_j)\,\Delta t_{\text{ex}} \tag{11}$$

which means the queue discharge rate q_d is:

$$q_d = \frac{n-1}{H} = \frac{(N-1)v_f}{s_{1,N}} = \frac{v_f \rho_{\text{cri}}}{1 + \rho_{\text{cri}}\,(v_j - v_f)\,\Delta t_{\text{ex}}} \tag{12}$$

3.2.2. Analysis of model properties The independent impact of the reaction time extension is evalued with Equation (12), see Figure 5(a). We examine the relation between the speed in cognestion and the queue discharge rate, setting reaction time extension Δt_{ex} to 0.05s, 0.1s, 0.15s and 0.2s. Figure 5(a) firstly indicates that reaction time extension Δt_{ex} can give a positive relation between the speed in congestion and the queue discharge rate. As the reaction time extension increases even slightly, the queue discharge rate will decrease considerably. When $\Delta t_{\text{ex}} = 0$s, the queue discharge rate equals to the capacity. Secondly, a dynamic reaction time extension can model the empirical observation. The bold line in Figure 5(a) is the empirical relation revealed in Yuan et al. (2015) (see Figure 1). The intersections between the bold line and the other lines indicates that to give empirical observations we may need to decrease the reaction time extension as the speed in congestion increases. Note that in Yuan et al. (2015) (see Figure 1) when the vehicular



(a) Sensitivity of the analytical model to the reaction time extension.

(b) An intra-driver reaction time extension mechanism for giving empirical observations.

Figure 5 Sensitivity of queue discharge rates to reaction time extensions.

speed in congestion reached around 63km/h, the queue discharge rate has reached 6840 veh/h which is much higher than a three-lane dutch freeway capacity 6300veh/h (with 15% proposition of trucks) claimed in Goemans (2011). Be more specific, that 6840 veh/h is the highest point measured in the data. We believe 6840veh/h is close to a three-lane freeway capacity. So it is believed that when the vehicular speed in congestion exceeds 63km/h, there is no capacity drop. We use $v_j^{\max}$ to indicate the lowest speed in congestion leading to no capacity drop. When $v_j \geqslant v_j^{\max}$, the queue discharge rate equals to the capacity. No capacity drop means the reaction time extension might be zero. Hence, we set

$$\Delta t_{\text{ex}} = \max\left(0, \gamma - \frac{\gamma v_j}{v_j^{\max}}\right) \tag{13}$$

γ is a parameter indicating the reaction time extension when the speed in congestion is 0 km/h. Varying with γ, we find a good relationship if we set $\gamma = 0.195$s. The modelled relation with Equation (13) is shown as dark triangulars in Figure 5(b). The bold line is the empirical relation as in Figure 1. The modelled relation can fit the empirical relation quite well, see Figure 5(b).

4. Numerical experiments.

In this section, we use numerical experiments to firstly validate the analytical model presented in Section 3.1.1. The estimation of queue discharge rate in Section 3.1.1 is an approximation with the Delta method. So we need to check whether the approximation is accurate enough. This validation step aims to make our conclusions solid.

Secondly, we present the combination effects of bounded acceleration spread and the reaction time extensions. A positive reaction time extension can allow a following vehicle to have a faster-than-predecessor acceleration. So the acceleration of the last vehicle in the queue a_N does not have

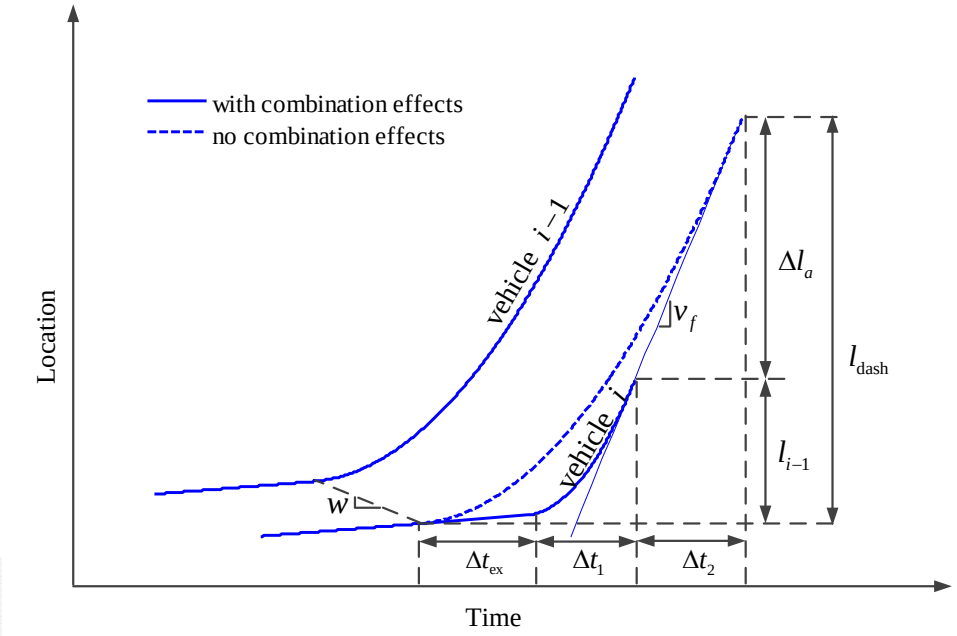


Figure 6 **Measurement of accelerations when reaction time is extended.**

to be the slowest acceleraion among all vehicles in the queue. The distribution of the last vehicles acceleration is difficult to deduce, so we decide to use numerical experiments to see the combination effects of bounded acceleration spread and the reaction time extension.

Thirdly, we try to see how to give a same relation between the speed in congestion and the queue discharge rate as empirical observations, considering combination effects of the acceleration spread and the reaction time extension.

The above three key questions are answered in three experiments denoted by Ex1, Ex2 and Ex3, respectively. The simulation results in this section correlate quite well with our analytical findings in Section 3. No matter whether the reaction time is included or not, the acceleration spread does not contribute sufficiently to the capacity drop. No matter whether the acceleration spread is considered, a negative relation between the reaction time and the speed in congestion can give similar queue discharge rates as empirical observations.

4.1. Simulation model.

Figure 6 shows trajectories of two vehicles accelerating from congestion, $i-1$ and i. Let us set an acceleration difference Δa. The free-flow spacing between vehicle i and $i-1$ will be $\frac{1}{\rho_{\text{cri}}}$ if $a_i = a_{i-1} + \Delta a$. In a special case that reaction time of all vehicles are not extended (i.e., $t_r = \Delta t$), $\Delta a = 0\text{m/s}^2$. So if $a_i > a_{i-1} + \Delta a$, the free-flow spacing between two vehicles will be smaller than the critical spacing $\frac{1}{\rho_{\text{cri}}}$. In Figure 6 we use a dashed line to present a trajectory of vehicle i according

to Newells model. Finally the trajectory of vehicle i will overlap with the dashed line. Vehicle i reached the free-flow speed earlier than the dashed trajectory by Δt_2. The whole acceleration process of vehicle i lasts Δt_1. So we can describe v_f as follows:

$$v_f = v_j + (a_{i-1} + \Delta a)\,\Delta t_1 \tag{14}$$

Meanwhile, vehicle $i-1$ spends $\Delta t_{\text{ex}} + \Delta t_1 + \Delta t_2$ accelerating up to v_f, i.e.,

$$v_f = v_j + a_{i-1}\,(\Delta t_{\text{ex}} + \Delta t_1 + \Delta t_2) \tag{15}$$

For the whole acceleration process, vehicle i travels $l_{i-1} = v_j \Delta t_{\text{ex}} + \frac{v_f^2 - v_j^2}{2\left(a_{i-1}+\Delta a\right)}$ which is less than the distance ($l_{\text{dash}} = \frac{v_f^2 - v_j^2}{2a_{i-1}}$) the dashed trajectory travel by $\Delta l_a = l_{\text{dash}} - l_{i-1}$. Vehicle i spends Δt_2 traveling Δl_a at free-flow speed v_f:

$$v_f = \frac{\frac{v_f^2 - v_j^2}{2a_{i-1}} - \left(v_j \Delta t_{\text{ex}} + \frac{v_f^2 - v_j^2}{2\left(a_{i-1}+\Delta a\right)}\right)}{\Delta t_2} \tag{16}$$

Combination of Equation (14),(15) and (16) can give:

$$\Delta a = \frac{2a_{i-1}^2 \Delta t_{\text{ex}}}{v_f - v_j - 2a_{i-1}\Delta t_{\text{ex}}}, \quad \text{for} \quad v_f - v_j > 2a_{i-1}\Delta t_{\text{ex}} \tag{17}$$

Equation (17) shows that the following vehicle can catch up with its predecessor with $a_i = a_{i-1} + \Delta a$ when $v_f - v_j > 2a_{i-1}\Delta t_{\text{ex}}$. If $a_{i-1} + \Delta a \leqslant a_i^{\text{desire}}$, then $a_i = a_{i-1} + \Delta a$, the free-flow spacing between vehicle i and $i-1$ will be critical spacing. If $a_{i-1} + \Delta a > a_i^{\text{desire}}$, then $a_i = a_i^{\text{desire}} < a_{i-1} + \Delta a$, the free-flow spacing between two successive vehicles is $s_i = \frac{1}{\rho_{cri}} + \frac{\left(\frac{1}{a_i} - \frac{1}{a_{i-1}}\right)\left(v_j - v_f\right)^2}{2} + (v_f - v_j)\,\Delta t_{\text{ex}}$.

When $v_f - v_j \leqslant 2a_{i-1}\Delta t_{\text{ex}}$, i.e., the reaction time is too long, and it is impossible for the follower to catch up with the leader. In this case, the followers acceleration will not be limited by its predecessor, i.e., $a_i = a_i^{\text{desire}}$. The free-flow spacing between two vehicles will be larger than the critical spacing, calculated as $s_i = \frac{1}{\rho_{\text{cri}}} + \frac{\left(\frac{1}{a_i} - \frac{1}{a_{i-1}}\right)\left(v_j - v_f\right)^2}{2} + (v_f - v_j)\,\Delta t_{\text{ex}}$. In summary, each vehicle accelerates with acceleration:

$$a_i = \min\left(a_{i-1} + \Delta a, a_i^{\text{desire}}\right) \tag{18}$$

and the free-flow spacing between any vehicle and its predecessor is:

$$s_i = \begin{cases} \frac{1}{\rho_{\text{cri}}}, & \text{for } v_f - v_j > 2a_{i-1}\Delta t_{\text{ex}} \text{ and } a_{i-1} + \Delta a \leqslant a_i^{\text{desire}} \\ \frac{1}{\rho_{\text{cri}}} + \frac{\left(\frac{1}{a_i} - \frac{1}{a_{i-1}}\right)\left(v_j - v_f\right)^2}{2} + (v_f - v_j)\,\Delta t_{\text{ex}}, & \text{for others} \end{cases} \tag{19}$$

In the numerical experiment we calculate the queue discharge rate as:

$$q_d = \frac{v_f}{E\,(s_i)}, \quad i = 2, ..., n \tag{20}$$

Equation (19) and (20) are general expressions for estimating queue discharge rates in three experiments, that is for the validation of analytical models (in Ex1), the examination of combination effects (in Ex2) and the reproduction of empirical observations (in Ex3), respectively.

Since in Section 3.1, we found the independent impact of acceleration on the queue discharge rate is marginal. We hypothesize that when considering reaction time extensions, the acceleration spread cannot contribute to queue discharge rate reduction greatly, either. The consequence of the hypothesis is that to obtain the empirically observed queue discharge rate (Figure 1), it is more important to model the impact of the reaction time extension than that of the acceleration spread. Hence, we still use Equation (13) to give the queue discharge rate in the third experiment Ex3.

4.2. Simulation set-up.

This section describes the set-ups of the three experiments. For validations of the analytical model in Section 3.2.1, we let $\Delta t_{\text{ex}} = 0$s in Ex1. For examining combination effects of the acceleration spread and the reaction time extension (Ex2), we set two scenarios, i.e., $\Delta t_{\text{ex}} = 0.1$s and $\Delta t_{\text{ex}} = 0.2$s. Finally, in the third experiment (Ex3) we set $\gamma = 0.18$s.

At the beginning of each experiment, we set all vehicles desired acceleration and reaction time extension. With Equation (18), (19) and (20), we can directly have the final queue discharge rate. The notations of models and the set-up of fundamental diagram are the same as those in Section 3. To draw the relation between the speed in congestion and the queue discharge rate, in each scenario set-up we run one simulation with newly distributed desired accelerations for each speed in congestion. We run the simulation for 1000 times to get the expected value and standard spread of queue discharge rates. We use $N = 660$veh, $a_{\min} = 0.5\text{m/s}^2$ and $a_{\max} = 2\text{m/s}^2$.

4.3. Validations of analytical models.

We approximate the mean queue discharge rate by approximating the expected value of the time-headway in Section 3.1. So we need to check whether the approximations are accurate enough to draw conclusions on the independent impacts of accelerations, which is finished in the first experiment Ex1. The comparison between the numerical experiment result and the analytical result is shown in Figure 7. In Figure 7, we use error bars and plus signs to indicate the standard spread and the expected value of queue discharge rates respectively for experiment results. Circles show the analytical approximations of queue discharge rates from Section 3.1.1.

We find that the analytical approximations of queue discharge rates fit the numerical experiment results well. Secondly, the queue discharge rate spread increases as the speed in congestion decreases. The fluctuation of queue discharge rate might be a related to the order of desired accelerations. More importantly, the analysis of the independent impacts of accelerations on the queue discharge rate in Section 3.1 is correct, that is, the acceleration spread cannot contribute sufficiently to the capacity drop independently.

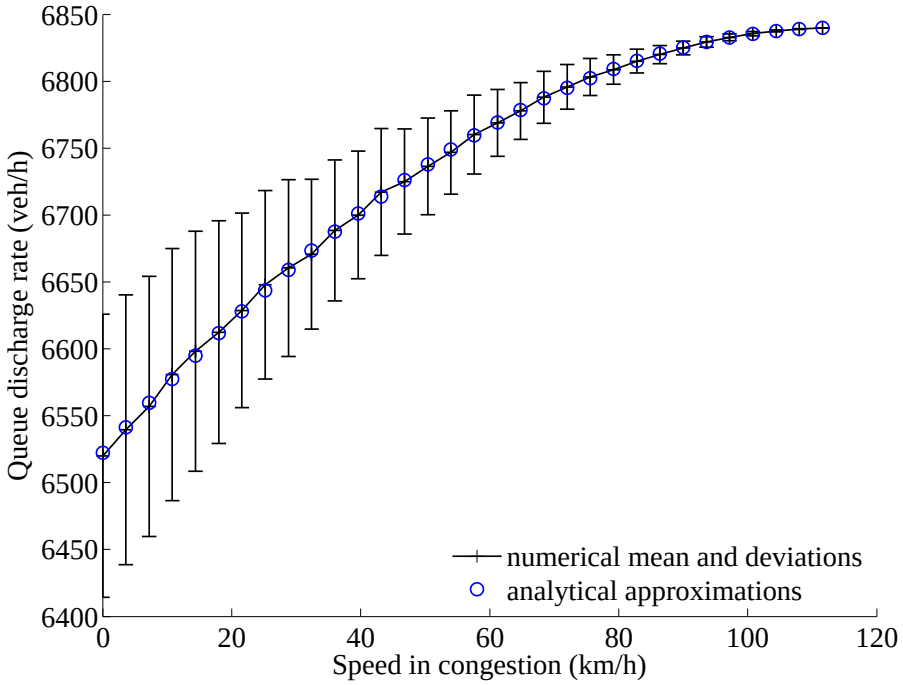


Figure 7 **Validation of the analytical model for the inter-driver acceleration spread. Mean and standard spreads of queue discharge rates are shown as plus signs and error bars, respectively**

4.4. Combination effects of the accelerations spread and reaction time extension.

The combination effects of the acceleration spread and the reaction time extension is examined in numerical experiments (Ex2), shown in Figure 8. As a reference, we use circles to indicate the mean queue discharge rate with the independent impact of reaction time, which is the same as shown in Figure 5(a).

Figure 8 shows that the acceleration spread hardly contributes to the queue discharge rate reduction. When setting $\Delta t_{\text{ex}} = 0.1$s and $v_j = 0$km/h, we find the acceleration spread can only reduce the queue discharge rate by 180 veh/h (around 3% reduction) at most. That observation is shown as the small difference between the plus signs and the circles in Figure 8. Meanwhile, increasing Δt_{ex} from 0.1s to 0.2s decreases the queue discharge rate considerably by 13% (with acceleration spread) and 14% (without acceleration spread) in a case of $v_j = 0$km/h. It also means a slight decrease of reaction time can contribute a considerable increase of queue discharge rates.

Because the acceleration spread can only reduce the queue discharge rate slightly, we use speed-dependent reaction time extension (see Equation (13)) to model mechanism of capacity drop to give queue discharge rates (Ex3). Figure 9 shows the experiment results. As reaction time decreases when congestion gets lighter, queue discharge rates can fit empirical observations well.

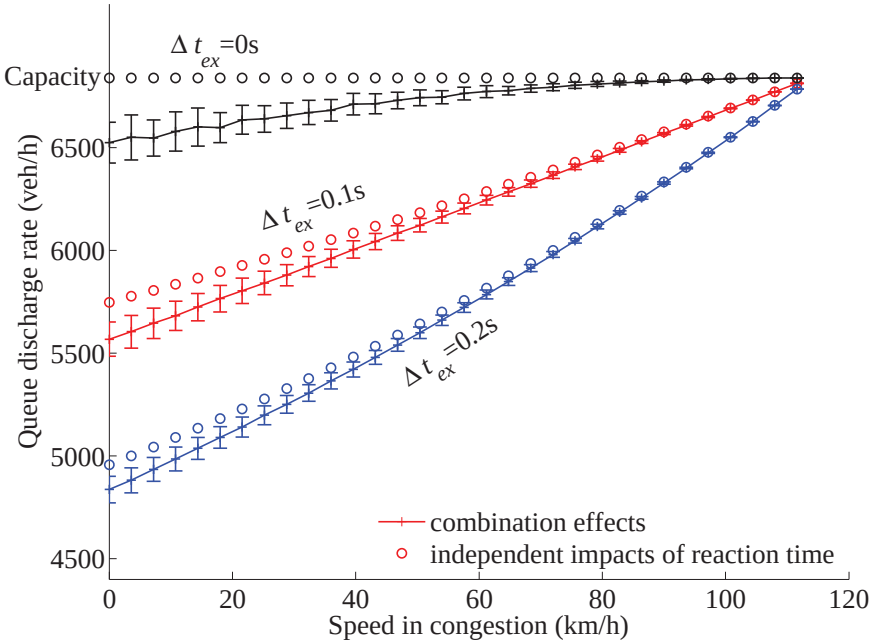


Figure 8 **Combination effects of inter-driver acceleration spread and reaction time extension on queue discharge rates.**

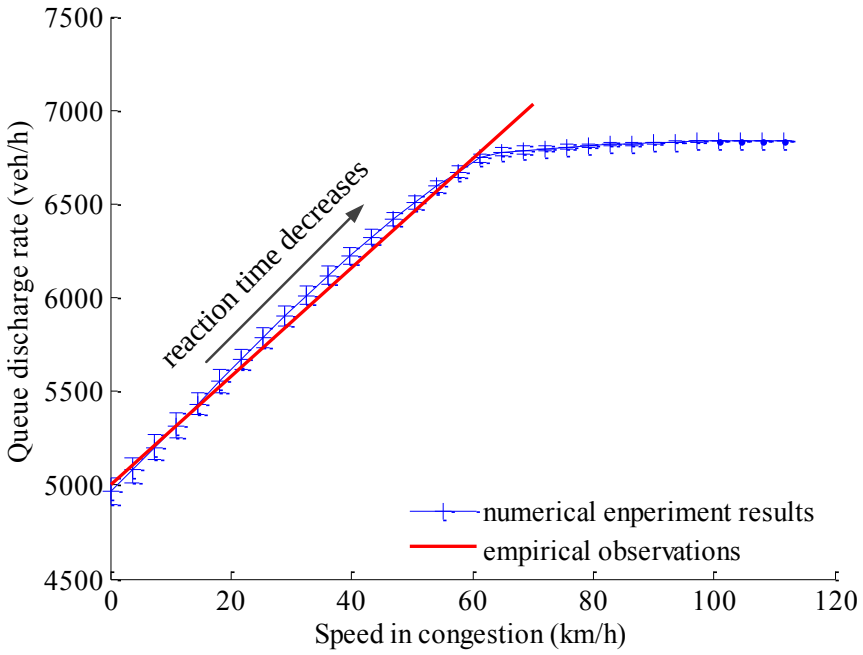


Figure 9 **Combination effects on queue discharge rates with intra-driver reaction time extension mechanism.**

5. Discussions and Conclusions.

This paper, for the first time as far as authors know, reveals that the impact of inter-driver acceleration spread on the queue discharge rate is rather magnitude. No matter whether the reaction time is considered or not, the inter-driver acceleration spread does not decrease the queue discharge rate as much as found empirically. On the contrary, intra-driver variety mechanism, that is the reaction time decreases as the speed in congestion increases, yields a similar relation between the speed in congestion and the queue discharge rate as found in empirical observations.

This research assumes the free-flow branch in the flow-density fundamental diagram as a straight line. Some researchers might argue that the free-flow speed in the vicinity of capacity is lower than the maximum free-flow speed, according to empirical observations. Because in empirical observations, the average free-flow speed in the vicinity of capacity is not very sensitive to the average density across lanes Yuan et al. (2016), relaxing the linear free-flow branch assumption by letting v_f as the free-flow speed in the vicinity of capacity and the maximum free-flow speed as v_max ($v_{\max} > v_f$), does not influence our research results at all.

Therefore, we conclude that including the acceleration spread when modeling the capacity drop within car-following models is not essential, but including reaction time variations is. Also, this paper gives reasons to believe that a control approach motivating drivers accelerate earlier might be able to considerably benefit maximizing queue discharge rates.

Acknowledgments

This research is financially supported by China Scholarship Council (CSC) and the NWO grant "There is plenty of room in the other lane". We thank anonymous referees for their helpful comments.

References

Banks, H, James. 1991. The two-capacity phenomenon: Some theoretical issues. *Transportation Research Record* (1320) p. 234–241–.

Bertini, Robert L., Monica T. Leal. 2005. Empirical study of traffic features at a freeway lane drop. *Journal of Transportation Engineering* **131**(6) 397–407. doi:10.1061/(ASCE)0733-947X(2005)131:6(397). URL `http://dx.doi.org/10.1061/(ASCE)0733-947X(2005)131:6(397)`.

Cassidy, Michael J, Robert L Bertini. 1999. Some traffic features at freeway bottlenecks. *Transportation Research Part B: Methodological* **33**(1) 25 – 42.

Cassidy, Michael J., Jittichai Rudjanakanoknad. 2005. Increasing the capacity of an isolated merge by metering its on-ramp. *Transportation Research Part B: Methodological* **39**(10) 896 – 913.

Chen, Danjue, Soyoung Ahn, Jorge Laval, Zuduo Zheng. 2014. On the periodicity of traffic oscillations and capacity drop: The role of driver characteristics. *Transportation Research Part B: Methodological* **59** 117 – 136.

Chen, Danjue, Jorge A. Laval, Soyoung Ahn, Zuduo Zheng. 2012. Microscopic traffic hysteresis in traffic oscillations: A behavioral perspective. *Transportation Research Part B: Methodological* **46**(10) 1440 – 1453.

Chung, Koohong, Jittichai Rudjanakanoknad, Michael J. Cassidy. 2007. Relation between traffic density and capacity drop at three freeway bottlenecks. *Transportation Research Part B: Methodological* **41**(1) 82 – 95.

Coifman, Benjamin, Seoungbum Kim. 2011. Extended bottlenecks, the fundamental relationship, and capacity drop on freeways. *Procedia - Social and Behavioral Sciences* **17**(0) 44–57.

Daganzo, Carlos F. 2005. A variational formulation of kinematic waves: basic theory and complex boundary conditions. *Transportation Research Part B: Methodological* **39**(2) 187–196.

Duret, Aurèlien, Jacques Bouffier, Christine Buisson. 2010. Onset of congestion from low-speed merging maneuvers within free-flow traffic stream. *Transportation Research Record: Journal of the Transportation Research Board* **2188** 96–107.

Goemans, W. Heikoop H., J.W. Daamen. 2011. *Dutch Highway Capacity Manual ("Handboek Capaciteitswaarden Infrastructuur Autosnelwegen (CIA) Volledig Vernieuwd").*

Hall, Fred L., Kwaku Agyemang-Duah. 1991. Freeway capacity drop and the definition of capacity. *Transportation Research Record* (1320) p. 91–98.

Kerner, B. S. 1998. Experimental features of self-organization in traffic flow. *Phys. Rev. Lett.* **81** 3797–3800.

Koutsopoulos, Haris N., Haneen Farah. 2012. Latent class model for car following behavior. *Transportation Research Part B: Methodological* **46**(5) 563–578.

Laval, Jorge A., Carlos F. Daganzo. 2006. Lane-changing in traffic streams. *Transportation Research Part B: Methodological* **40**(3) 251 – 264.

Lebacque, J. 2003. Two-phase bounded-acceleration traffic flow model: Analytical solutions and applications. *Transportation Research Record: Journal of the Transportation Research Board* **1852** 220–230.

Leclercq, Ludovic, Victor L. Knoop, Florian Marczak, Serge P. Hoogendoorn. 2015. Capacity drops at merges: New analytical investigations. *Transportation Research Part C: Emerging Technologies –.*

Leclercq, Ludovic, Jorge A. Laval, Nicolas Chiabaut. 2011. Capacity drops at merges: An endogenous model. *Transportation Research Part B: Methodological* **45**(9) 1302 – 1313.

Lighthill, M. J., G. B. Whitham. 1955. On kinematic waves. ii. a theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **229**(1178) 317–345.

Newell, G.F. 2002. A simplified car-following theory: a lower order model. *Transportation Research Part B: Methodological* **36**(3) 195–205.

Nishinari, Katsuhiro, Martin Treiber, Dirk Helbing. 2003. Interpreting the wide scattering of synchronized traffic data by time gap statistics. *Phys. Rev. E* **68**(6) 067101.

Oh, Simon, Hwasoo Yeo. 2015. Impact of stop-and-go waves and lane changes on discharge rate in recovery flow. *Transportation Research Part B: Methodological* **77** 88 – 102.

Papageorgiou, M., I. Papamichail, A.D. Spiliopoulou, A.F. Lentzakis. 2008. Real-time merging traffic control with applications to toll plaza and work zone management. *Transportation Research Part C: Emerging Technologies* **16**(5) 535–553.

Reiss, Rolf-Dieter. 1989. *Approximate Distributions of Order Statistics: with applications to nonparametric statistics*. Springer-Verlag.

Richards, Paul I. 1956. Shock waves on the highway. *Operations Research* **4**(1) 42–51.

Srivastava, Anupam, Nikolas Geroliminis. 2013. Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. *Transportation Research Part C: Emerging Technologies* **30** 161 – 177.

Tampère, Chris. 2004. Human-kinetic multiclass traffic flow theory and modelling with application to advanced driver assistance systems in congestion. Ph.D. thesis, Delft University of Technology.

Treiber, Martin, Arne Kesting, Dirk Helbing. 2006. Understanding widely scattered traffic flows, the capacity drop, and platoons as effects of variance-driven time gaps. *Phys. Rev. E* **74**(1) 016123–.

Wong, GCK, SC Wong. 2002. A multi-class traffic flow modelan extension of lwr model with heterogeneous drivers. *Transportation Research Part A: Policy and Practice* **36**(9) 827–841.

Yeo, Hwasoo. 2008. Asymmetric microscopic driving behavior theory. Ph.D. thesis, University of California Transportation Center. UC Berkeley: University of California Transportation Center.

Yuan, Kai, Victor L. Knoop, Serge P. Hoogendoorn. 2015. Capapcity drop: relation between speed in congesiton and the queue discharge rate. *Transportation Research Record: Journal of the Transportation Research Board* **2491** 72–80.

Yuan, Kai, Victor L. Knoop, Ludovic Leclercq, Serge P. Hoogendoorn. 2016. Capacity drop: a comparison between stop-and-go wave and standing queue at lane-drop bottleneck. *Transportmetrica B: Transport Dynamics* (0) 1–14. doi:10.1080/21680566.2016.1245163. URL `http://dx.doi.org/10.1080/21680566.2016.1245163`.

Zhang, H.M., T. Kim. 2005. A car-following theory for multiphase vehicular traffic flow. *Transportation Research Part B: Methodological* **39**(5) 385–399.