Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 38 (2019) 668–687

Transportation Research Procedia

www.elsevier.com/locate/procedia

23rd International Symposium on Transportation and Traffic Theory, ISTTT 23, 24-26 July 2019, Lausanne, Switzerland

Continuum car-following model of capacity drop at sag and tunnel bottlenecks

Kentaro Wada[a], Irene Martínez[b], Wen-Long Jin[b,*]

[a]*Institute of Industrial Science, The University of Tokyo, Komaba 4-6-1, Meguro-ku, Tokyo 153-8505, Japan*
[b]*Institute of Transportation Studies, University of California, Irvine, 4000 AIR Bldg, Irvine, CA 92697, United States*

Abstract

Sags and tunnels are major bottlenecks, where the road capacity is reduced, and the "capacity drop" phenomenon occurs; however, there is no simple model or theory that can explain the formation and other characteristics of capacity drop. This paper presents a second-order car-following model, which is equivalent to a continuum model in the Lagrangian coordinates. The model is built on two main assumptions: (i) inhomogeneous fundamental diagrams with location-dependent time gaps, and (ii) bounded acceleration. We first demonstrate that the stationary speed profiles, the low acceleration rates, the dropped capacity, and the degree of the time duration of the capacity drop formation in the model are consistent with empirical observations. We then investigate the impacts on the stationary states and dropped capacity of the numerical viscosity caused by the discretization method and show that the dropped capacity converges to the theoretical value. We further propose a one-dimensional iterated function system model for the formation mechanism of the capacity drop, which is derived by investigating the spatial pattern of equilibrium and bounded acceleration traffic states that arises in a lead-vehicle problem. Utilizing this model, we uncover a set of properties of the capacity drop such as existence, uniqueness, global convergence, and convergence speed. We finally apply the model to analyze the impacts of heterogeneous drivers. The model and insights in this study will help to develop control and management schemes to alleviate capacity drop effects with connected and autonomous vehicles in the future.

Peer-review under responsibility of the scientific committee of the 23rd International Symposium on Transportation and Traffic Theory.

Keywords: Sags and tunnels; Capacity drop; Bounded acceleration; Car-following model; Iterated function system; Heterogeneous drivers

1. Introduction

Sags and tunnels are traffic bottlenecks where the road capacity can be substantially reduced compared with a normal section ("capacity reduction"); the so-called "capacity drop" further occurs once an upstream queue forms. Such bottlenecks have been recognized for many years (Edie and Foote, 1958; Koshi et al., 1992), but it is still one of the most common causes of congestion, especially in mountainous areas. On intercity expressways in Japan nearly 80% of traffic congestion occurs at sags or uphills (60%) and tunnel entrances (20%) (Xing et al., 2014).

* Corresponding author. Tel.: +1-949-824-1672 ; fax: +1-949-824-8385.
E-mail address: wadaken@iis.u-tokyo.ac.jp (K. Wada), wjin@uci.edu (W.-L. Jin)

2352-1465
Peer-review under responsibility of the scientific committee of the 23rd International Symposium on Transportation and Traffic Theory.
10.1016/j.trpro.2019.05.035

Koshi et al. (1992) and Koshi (2003) suggested that the above macroscopic phenomena are related to the following microscopic behaviors: (i) the congestion is triggered by lane-changing and/or particular car-following behaviors (i.e., compared to a normal section, a longer spacing is required and/or the free-flow speed tends to decrease, on average) and (ii) the capacity drop (or the queue discharge flow rate) is characterized by a low acceleration rate from the head of the upstream queue. Several empirical studies have supported the existence of these microscopic behaviors. Patire and Cassidy (2011) reported the roles of lane-changing (and the resulting speed disturbances) in the occurrence of congestion in detail. Goñi Ros et al. (2013) showed that average headways at similar speeds before the occurrence of persistent congestion increase inside a sag section. Similar observations were made by Koshi (2003) with a real tunnel section data and by Yoshizawa et al. (2012) with a driving simulation experiment at hypothetical sag sections. Furuichi et al. (2003) and Brilon and Bressler (2004) examined the reduction in the free-flow speed at sag and uphill sections. For the suggestion (ii), several studies showed that it typically takes 1–3 kilometers to accelerate from the speed at the upstream queue to the free-flow speed (Koshi et al., 1992; Okamura et al., 2000a; Ozaki, 2003). Moreover, lane changes are not necessary conditions for capacity reduction or capacity drop because these are observed even on a single-lane road (Okamura et al., 2000b).

Several car-following models have been developed to simulate traffic dynamics at such bottlenecks. These models are mainly intended to reproduce the capacity reduction phenomenon by describing the suggested behavior (i) in association with the road (grade and light) conditions. Koshi et al. (1992) developed a model that combines the General Motors type car-following model with the effect of a change in the vertical gradient in the sag section. This model was refined by Ozaki (1994), Xing and Koshi (1995) and Oguchi and Konuma (2009). Similar models were proposed by Komada et al. (2009) and Goñi Ros et al. (2016) based on the optimal velocity model (Bando et al., 1995) and the intelligent driver model (Treiber et al., 2000), respectively.

The above models may simulate the occurrence of certain congestion events by calibrating the model parameters appropriately. However, they have two main drawbacks. First, the models are too complicated for their physical/mathematical properties to be investigated. This causes difficulty in obtaining a systematic understanding of the relationship between microscopic behavioral rules and macroscopic phenomena. Second, the target phenomenon is also complicated, i.e., the capacity reduction rate is generally uncertain and widely distributed (Okamura et al., 2000b; Oguchi, 2004). This makes the validation of the models difficult. Note that, in contrast, traffic states such as the flow, speed, and also very low acceleration process during persistent congestion are almost stationary[1] and stable (Koshi et al., 1992; Ozaki, 2003) (we refer to this state as "capacity drop stationary state"). Unfortunately, however, the existing studies paid less attention to the stable phenomenon associated with the capacity drop, and whether the existing models reproduce it properly is not well explored. One exception is the model of Ozaki (1994, 2003), but it requires many (i.e., more than 20) parameters and is difficult to validate.

As the first step toward a simple model to describe the above bottleneck phenomena, Jin (2018) proposed a first-order kinematic wave (KW) model for sag and tunnel bottlenecks. Assuming location-dependent time gaps inside the bottleneck, the model describes the capacity reduction effect. Imposing a bounded acceleration constraint on the stationary states inside the capacity reduction zone, the model reproduces the occurrence of the capacity drop. It also captures the associated low acceleration rate when vehicles accelerate away from the upstream queue. Further, the model was verified with empirical observations of stationary vehicles' speed profiles at the Kobotoke tunnel in Japan (Koshi et al., 1992). Thus, despite its inability to describe rather complicated phenomena such as the (probabilistic) capacity reduction and stop-and-go traffic upstream of bottlenecks (Goñi Ros et al., 2014) at the current stage, this simplified model can be a suitable building block for a more comprehensive theory.

An important limitation of the KW model is that the capacity drop stationary states are assumed to be reached *instantaneously* once congestion occurs. However, the empirical studies indicate that there is a transition period to reach these states, which can vary from several minutes to one hour (Koshi et al., 1992; Ozaki, 2003; Moriyama et al., 2011)[2]. The main issue here is not how realistic the assumption is, but that the KW model cannot explain how the stationary state is formed and stabilized (i.e., second-order or dynamical properties of the capacity drop).

[1] In this study "stationary" means (only) time-independent, i.e., stationary states are time-independent but can be location-dependent.

[2] Koshi et al. (1992) stated that transition time (*Tt* in Figures 1 and 2 of the reference) is approximately an hour. Figures 5 and 7 of Ozaki (2003) and Figure 4 of Moriyama et al. (2011) indicated a relatively short transition time (i.e., 15 - 30 min.).

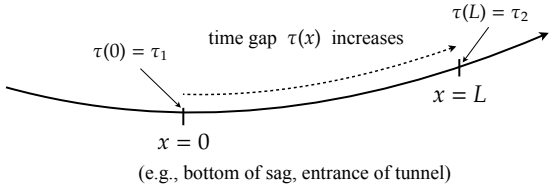


Fig. 1: Illustration of a sag or tunnel bottleneck

Furthermore, understanding these mechanisms is crucial to design an adequate and effective control to prevent (or mitigate) the capacity drop.

This paper presents a second-order car-following model, which is equivalent to a continuum model in the Lagrangian coordinates, for the capacity drop at sag and tunnel bottlenecks. The model is built on the same two assumptions as Jin (2018): (i) inhomogeneous fundamental diagrams (FDs) with location-dependent time gaps, and (ii) bounded acceleration (BA). We first demonstrate that the stationary states, dropped capacity and other characteristics are consistent with empirical results. The consistency with the KW model (Jin, 2018) at the capacity drop stationary states is also discussed through analyzing the effects of a numerical viscosity on the capacity drop; we observe that a larger numerical viscosity leads to a smaller dropped capacity. Then, we propose a one-dimensional iterated function system model for the formation mechanism of the capacity drop, which is derived by investigating the spatial pattern of equilibrium and bounded acceleration traffic states that arises in a lead-vehicle problem. Utilizing this model, we uncover a set of properties of the capacity drop such as existence, uniqueness, global convergence, and transition time. We finally apply the model to analyze the impacts of heterogeneous drivers for testing the robustness of the (homogeneous) model and obtaining a new theoretical prediction. We find that, on average, the first-order and second-order properties preserve when the FD is heterogeneous among drivers but that heterogeneous acceleration bounds lead to a systematic decrease in the dropped capacity.

The rest of the paper is organized as follows. Section 2 presents a second-order car-following model based on a unified framework for modeling BA by Jin and Laval (2018). Even though the proposed model is a direct extension of that in Jin and Laval (2018), by incorporating a location-dependent time gap, the model has fundamentally different behaviors as it reproduces the capacity drop phenomenon, which was shown impossible with a location-dependent speed limit. In addition, the ensuing analyses are all new. With a numerical example, we demonstrate the occurrence of the capacity drop and compare its characteristics to the known empirical observations. Section 3 analyzes the effects of a numerical viscosity on the capacity drop and discusses the consistency with the KW model by Jin (2018). Section 4 analyzes the formation and stabilized mechanisms of the capacity drop. Section 5 extends the analysis of the previous section to cases with heterogeneous drivers. Section 6 summarizes the study and discusses future directions.

2. A second-order model of capacity drop at sag/tunnel bottlenecks

2.1. Equivalent continuum and car-following models with inhomogeneous time gap and bounded acceleration

This section presents a second-order model for sag and tunnel bottlenecks illustrated in Fig. 1 based on a BA–LWR model (i.e., the LWR model with bounded acceleration) (Jin and Laval, 2018):

$$k_t + (kv)_x = 0, \tag{1a}$$

$$v_t + vv_x = \min\left\{A(x,v), \frac{V(x,k)-v}{\epsilon}\right\}, \tag{1b}$$

where k and v are the density and speed, respectively, and $V(x,k)$ is the speed-density FD. We denote the flow-rate by q. The subscript variables represent partial derivatives, and $\epsilon = \lim_{\Delta t\to 0+}\Delta t$ is a hyperreal infinitesimal number (for more detail, see Jin, 2017c; Jin and Laval, 2018).

The present model is built on the following two main assumptions. First, we use an inhomogeneous triangular FD

$$v = V(x,k) \equiv \min\left\{u, \frac{1}{\tau(x)}\left(\frac{1}{k}-\frac{1}{\kappa}\right)\right\}, \tag{2}$$

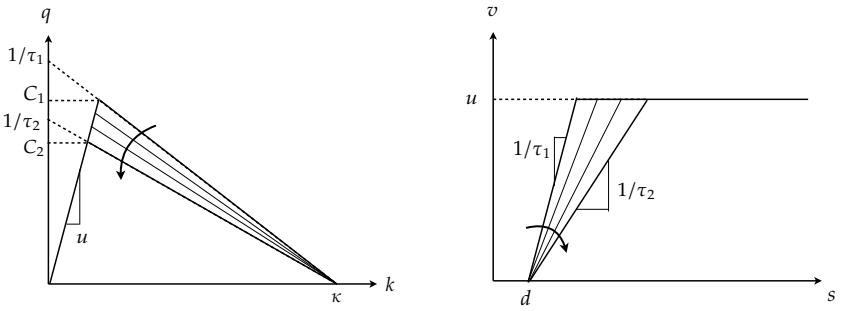


Fig. 2: Fundamental diagrams: flow-density relation (left); speed-spacing (right)

where u, κ, and $\tau(x)$ are the free-flow speed, jam density, and location-dependent time gaps, respectively. More specifically, we assume that $\tau(x)$ is an increasing function with x inside a bottleneck between $x = 0$ and $x = L$ (see Fig. 1) (w.l.o.g., $\tau(x) = \tau_1,\ x \notin [0, L]$), based on the observation that more work is needed on the throttle when accelerating on an upgrade or more careful driving under different light conditions inside the tunnel. This assumption is the same with that of the first order model by Jin (2018) and is consistent with the observations (Koshi et al., 1992; Koshi, 2003; Goñi Ros et al., 2013; Sun et al., 2018). The capacity also depends on the location

$$C(x) = \frac{u\kappa}{1 + u\kappa\tau(x)},$$

which decreases inside the bottleneck and reaches a minimum at $x = L$. We adopt the notations $C_1 = C(0)$, and $C_2 = C(L)$. Since $\tau_1 < \tau_2$, and $C_2 < C_1$, the inhomogeneous FD captures the capacity reduction phenomenon[3]. Note that L is one of the model parameters and has to be calibrated with empirical data (see Jin, 2018, and Section 2.2). However, the location at which congestion begins is generally 500-1000 m downstream of the bottom of sags (Brilon and Bressler, 2004; Hatakenaka et al., 2006) and a reasonable value of L would be slightly larger than 500-1000 m. The resulting flow-density FD, $q = Q(x, k) \equiv \min\left\{uk, \frac{1}{\tau(x)}\left(1 - \frac{k}{\kappa}\right)\right\}$, and the corresponding speed-spacing FD, which is used later, are shown in Fig. 2. Throughout this paper, we consider a linearly increasing time gap function:

$$\tau(x) = \tau_1 + \frac{x\Delta\tau}{L}, \quad \text{where } \Delta\tau \equiv \tau_2 - \tau_1 \qquad x \in [0, L]. \tag{3}$$

The second assumption limits the acceleration rate a $(= v_t + vv_x)$, which should be bounded due to limitations in the vehicles' mechanics, drivers' responses, road geometry etc. The BA function, $A(x, v)$, is assumed to have the following properties: (i) being non-negative (i.e., $A(x, v) \geq 0$); (ii) having a maximum start-up (i.e., $a_0 = A(x, 0)$); (iii) being bounded (i.e., $A(x, v) \leq a_0$); and (iv) being not increasing in speed (i.e., $A_v(x, v) \leq 0$). In this study we use the following constant (or state-independent) BA model,

$$A(x) = a_0 - g\Phi(x), \tag{4}$$

where a_0 is the maximum acceleration rate on a level road, $g = 9.8\ m/s^2$ is the acceleration due to gravity, and $\Phi(x)$ is the decimal grade at x. Another one widely in the literature is the TWOPAS BA model (Allen et al., 2000).

The model (1) is equivalent to the two-phase model proposed by Lebacque (2003): the traffic state is either in the equilibrium phase, if $v = V(x, k)$ and $a \leq A(x, v)$; or in the BA phase, if $v \leq V(x, k)$ and $a = A(x, v)$. Jin and Laval (2018) solved this model with location-dependent speed limits (or free-flow speeds) and showed that such an inhomogeneity does not cause the capacity drop. In this study, we are interested in solving the model (1) with the inhomogeneous FD (2) and examining whether it can describe the capacity drop phenomenon[4].

[3] The reduction in the free-flow speed in sag sections (Furuichi et al., 2003; Brilon and Bressler, 2004; Sun et al., 2018) may be another factor behind the capacity reduction. We, however, assume a constant free-flow speed because this factor does not affect the capacity drop phenomenon.

[4] In Jin (2018), traffic states inside the bottleneck and an auxiliary segment between $x = L$ and $x = L+L'$ are assumed to be always instantaneously stationary, i.e., $v_t = 0,\ x \in [0, L + L']$ in Eq.(1). Thus, the second-order properties of the capacity drop can not be described.

To do so, the model (1) is converted to a car-following model using the methodology developed by Jin (2016) as follows,

$$X(t+\Delta t,n) = X(t,n) + \min\left\{V(x,s(t,n)), v(t,n) + A(x,v)\Delta t\right\}\Delta t \tag{5a}$$

$$V(x,s(t,n)) = \min\left\{u, \frac{1}{\tau(X(t,n))}(s(t,n)-d)\right\}. \tag{5b}$$

where $X(t,n)$, $v(t,n)$, and $s(t,n) \equiv \frac{X(t,n-\Delta n)-X(t,n)}{\Delta n}$ are the position, the speed, and the spacing of n-th vehicle at time t, and $V(x,s)$ is the speed-spacing FD where d is the jam spacing. The model (5) with $\Delta n = 1$ can be viewed as an inhomogeneous generalization of Newell's car-following model (Newell, 2002) and its BA version (Laval and Leclercq, 2008). If $\Delta n \to 0$ for a fixed ratio $\frac{\Delta t}{\Delta n}$, the model (5) will converge to an equivalent (LWR-BA) continuum model in Lagrangian coordinates. Note that, when $\Delta n < 1$, the vehicles with non-integer numbers can be interpreted as "imaginary vehicles" or interpolations to ensure the proper information propagation speed and the reaction of real vehicles, i.e., with integer numbers[5]. We thus call the model a "continuum car-following model."

It is clear from (5) that vehicle n is affected only by the vehicle ahead, vehicle $n-\Delta n$, which is the *anisotropic property* (Daganzo, 1995). The car-following model also satisfies the *collision-free property* (Jin, 2016) when choosing appropriate discretized units Δt and Δn that satisfy

$$\max\left(\frac{\partial V(x,s)}{\partial s}\right) = \max\left(\frac{1}{\tau(X)}\right) = \frac{1}{\tau_1} \le \frac{\Delta n}{\Delta t}. \tag{6}$$

For triangular FDs, the condition (6) is equivalent to the Courant–Friedrichs–Lewy (CFL) condition of a Godunov finite difference equation for a hyperbolic conservation formulation of the LWR model in Lagrangian coordinates (Leclercq et al., 2007). In general, the collision-free property is stronger than CFL condition and should be applied, e.g. non-concave fundamental diagrams (Jin, 2017c). Moreover, when the collision-free condition is satisfied, the spacing never smaller than the jam spacing, d, and the vehicle's speed is thus non-negative, i.e., the car-following model satisfies the *forward traveling property*. Note that the deceleration rate in the model is not bounded, i.e., the deceleration behavior is exactly the same as the LWR model. Thus, one may consider the model as a one-and-half order model.

2.2. Occurrence and characteristics of capacity drop

To demonstrate the consistency with empirical observations, we will show a numerical example of a lead-vehicle problem with the car-following model (5), where the leader drives at the free-flow speed and the followers arrive at the entrance of the road section with a constant headway. We calibrated the model with the constant BA model (4) using the empirical observations of stationary speed profiles of four vehicles at the Kobotoke tunnel in Japan, which is also the upgrade section (Figure 3 of Koshi et al., 1992, and see also Fig. 3): $L = 1500$ [m], $u = 75$ [km/h], $\kappa = 140$ [veh/km], $\tau_1 = 1.5$ [sec], $\tau_2 = 2.1$ [sec], $a_0 = 0.312$ [m/s^2], $\Phi(x) = \{0.023,\ x \in [0,L];\ 0.023 - 0.063(x-1.5),\ x \in (L, L+1000);\ -0.04,\ x \ge L+1000\}$, $\Delta n = 0.05$, and $\Delta t = 0.075$ [sec][6]. In this setting, $C_1 = 1953.5$ [veh/h/lane], $C_2 = 1473.7$ [veh/h/lane]. Note that we "estimated" only L and a_0 (by hand) and other parameters were set as directly observable or typical values (see, Jin, 2018, for the calibration procedure). The theoretical dropped capacity, C^-, is 1324.7 [veh/h/lane], based on formula (Eq.(32) in Jin, 2018). Thus, the capacity drop ratio, $\zeta \equiv 1 - \frac{C^-}{C_2}$, is 0.10. The arrival flow rate at the road entrance is set as $1.05C_2$ (i.e., the demand exceeds the bottleneck capacity).

Fig. 3 shows the speed profile inside and downstream of the bottleneck predicted by the car-following model when flows inside the bottleneck become (almost) stationary (after t_s in Fig. 4(c)). Here $t_s \approx 20$ min, as defined in Section 4, whose degree is consistent with those of the observed transition times (Moriyama et al., 2011). Fig. 4 (a) shows the (real) vehicle trajectories where colors indicate traffic phases, Fig. 4 (b) shows the same vehicle trajectories inside

[5] If the car-following interpretation is taken when $\Delta n < 1$, it should be better to consider only the real vehicles to analyze the results (trajectories, flow rates, passing times, etc.), although both real and imaginary vehicles are considered in the simulation.

[6] The calibrated parameters are slightly different from those of the KW model with the TWOPAS model in Jin (2018), especially for the free-flow speed. Thus, the resulting C_1 and C_2 values are slightly lower.

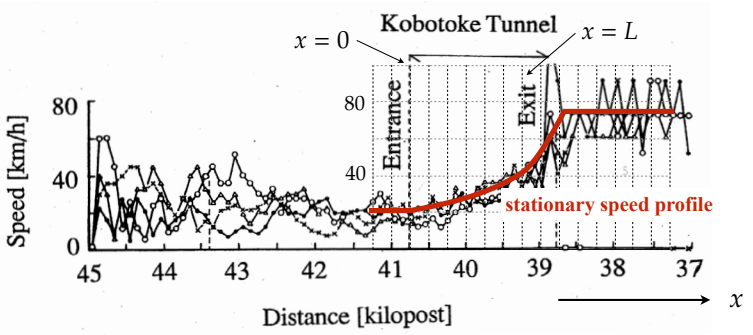


Fig. 3: Stationary speed profile inside and downstream of bottleneck (red line: proposed model; background: Fig. 3 of Koshi et al., 1992)

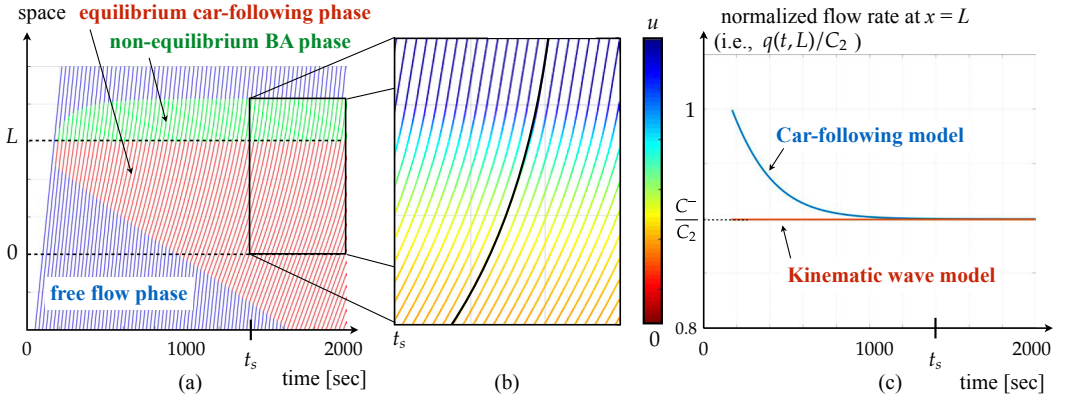


Fig. 4: Numerical example: (a) vehicle trajectories (where color indicates traffic phase); (b) vehicle trajectories inside standing waves (where color indicates speed); (c) evolution of queue discharge flow rate at $x = L$.

standing waves where colors indicate speeds, and Fig. 4 (c) shows the evolution of the normalized queue discharge flow rate (i.e., the reciprocal of the headway between two real vehicles) at the end of the bottleneck, $x = L$.

Fig. 3 demonstrates that the stationary speeds (i.e., the very low acceleration process) inside the bottleneck match the observation very well. Fig. 4(c) shows that the capacity drop occurs and the flow rate converges to C^-, which is a typical value of the dropped capacity at the same site (Koshi et al., 1992). The transition time from the free-flow to capacity drop stationary state is about 20 minutes, which is comparable with a typical transition time (30 minutes) reported in Japan Highway Public Corporation (1987) based on one-year observation of congestion at the same site. This may be somewhat surprising since we did not use any data during the *transition period* for the model calibration. The model also replicates a well-known feature of the (spatial) transition of the acceleration location inside the bottleneck (Koshi et al., 1992). That is, the location at which vehicles start to accelerate is finally stayed at the upstream end of the bottleneck $x = 0$, although the onset of congestion occurs around the downstream end of the bottleneck (Figs. 4(a) and (b)). The consistency with some empirical features above suggests that exploring the mechanism behind the characteristics of the second-order model can be helpful for understanding the capacity drop phenomenon at sags and tunnels. We will pursue this point in Sections 4 and 5.

3. Impact on stationary state and capacity drop of discretization method

The collision-free condition dictates that $\frac{\Delta t}{\tau_1 \Delta n} \leq 1$. In the above numerical example (see Fig. 4), the ratio of $\frac{\Delta t}{\Delta n} = \tau_1$ is assumed to be constant (i.e., the CFL condition is satisfied as an equality) and $\Delta t \to 0$. It is known that when modeling fluid flow, certain numerical approximations to discretize the model can introduce a numerical viscosity. Traditionally this is handled by repeating the simulation at different grid sizes to determine whether there is a grid size effect on the solutions. In this section, we will theoretically and numerically study how much numerical viscosity is

introduced once a discretization, Δt and Δn, is chosen. In particular, we analyze how such numerical viscosity affects the stationary states and the dropped capacity. These discussions are critical to understand the property of the new model, especially the difference and consistency between the model and the KW model.

3.1. Analytical results

We model the numerical viscosity by the following viscous LWR model:

$$k_t + (kV(x,k))_x = \alpha k_{xx}, \tag{7}$$

where α is a positive constant and k_{xx} is the second partial derivative of k. Effectively, we have an extended fundamental diagram:

$$q = \tilde{Q}(x,k,k_x) \equiv kV(x,k) - \alpha k_x, \quad v = \tilde{V}(x,k,k_x) \equiv V(x,k) - \alpha\frac{k_x}{k}. \tag{8}$$

In congested stationary states, the flow rate, q^*, is constant. The stationary density of the upstream link is $k^*(x) = (1-\tau_1 q^*)\kappa,\ x \le 0$. Inside the bottleneck ($x \in [0,L]$), the density is given by the ordinary differential equation:

$$q^* = \frac{1}{\tau(x)}\left(1 - \frac{k^*(x)}{\kappa}\right) - \alpha k_x^*(x), \tag{9}$$

where $k^*(0) = (1-\tau_1 q^*)\kappa$ [7]. The above equation is solved as

$$k^*(x;q^*) = \kappa - \kappa q^*\frac{\Delta\tau}{L}\frac{1}{1+\tilde{\alpha}}\left(x + \frac{\tau_1}{\Delta\tau}L\right) + \kappa\beta\left(1 + \frac{x}{L}\frac{\Delta\tau}{\tau_1}\right)^{-1/\tilde{\alpha}}, \tag{10}$$

where $\tilde{\alpha} = \alpha\frac{\kappa\Delta\tau}{L}$ is dimensionless and $\beta = -q^*\tau_1\frac{\tilde{\alpha}}{1+\tilde{\alpha}}$ is also dimensionless and can be obtained by substituting $k^*(0)$ into the equation.

When $\tilde{\alpha} \ll 1$, the stationary density inside the bottleneck can be approximated as

$$k^*(x;q^*) \approx \kappa\left[1 - \frac{\gamma}{L}\left(x + \frac{\tau_1}{\Delta\tau}L\right)\right], \tag{11}$$

where $\gamma = q^*\Delta\tau(1-\tilde{\alpha})$ is dimensionless. Thus, for $x \in [0,L]$, $k^*(x;q^*)$ decreases with increasing x; hence, $v^*(x;q^*) = q^*/k^*(x;q^*)$ increases with increasing x, and $a^*(x;q^*) = v^*(x;q^*)v_x^*(x;q^*)$ increases with increasing x.

Thus, if $A(x,v)$ is non-increasing for both increasing v and x, the adjusted dropped capacity, $\tilde{C}^-$, satisfies

$$A(L, v^*(L;\tilde{C}^-)) = a^*(L;\tilde{C}^-) \approx \frac{(\tilde{C}^-)^2\gamma}{L\kappa^2(1-\gamma\frac{\tau_2}{\Delta\tau})^3}, \tag{12}$$

where $\gamma = \tilde{C}^-\Delta\tau(1-\tilde{\alpha})$. The corresponding capacity drop ratio is $\tilde{\zeta} = 1 - \frac{\tilde{C}^-}{C_2}$.

Theorem 3.1. *For a constant acceleration bound, $A(x) = a_0 - g\Phi(x)$, a larger numerical viscosity, α, leads to a larger dropped capacity and a smaller capacity drop ratio.*

Proof. From the definition, we have

$$a^*(x;q^*) \approx \frac{(q^*)^2\gamma}{L\kappa^2[1-\gamma(\frac{x}{L} + \frac{\tau_1}{\Delta\tau})]^3}.$$

It can be shown that $a^*(x;q^*)$ increases with increasing q^* and decreases with increasing α. From (12), we can conclude that a larger α leads to a larger dropped capacity and a smaller capacity drop ratio. ■

[7] We here assume that $k^*(x)$ is continuous at $x = 0$.

Table 1: Dropped capacity $\tilde{C}^-$ [veh/h] and corresponding viscosity (α) [km^2/h] for different Δt and Δn in the numerical example.

$\frac{\Delta t}{\Delta n}$	$\Delta n = 1$	$\Delta n = 0.5$	$\Delta n = 0.25$	$\Delta n = 0.1$	$\Delta n = 0.02$
1.5	1337.7 (0.738)	1331.4 (0.382)	1328.0 (0.189)	1326.0 (0.075)	1324.9 (0.013)
1	1341.2 (0.934)	1333.1 (0.478)	1328.9 (0.240)	1326.3 (0.096)	1325.0 (0.017)
0.5	1344.6 (1.122)	1334.8 (0.575)	1329.8 (0.292)	1326.7 (0.116)	1325.0 (0.022)
0.25	1346.2 (1.212)	1335.7 (0.624)	1330.2 (0.316)	1326.9 (0.127)	1325.1 (0.024)
0.05	1347.5 (1.283)	1336.3 (0.662)	1330.6 (0.336)	1327.0 (0.135)	1325.1 (0.025)

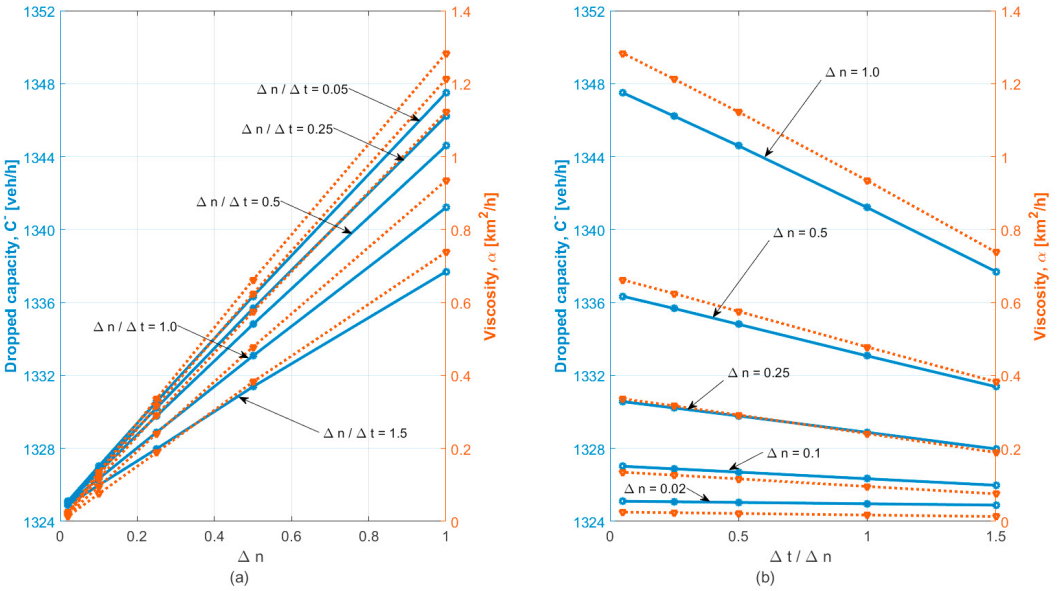


Fig. 5: Numerical dropped capacity, $\tilde{C}^-$, and numerical viscosity, α: (a) convergence of dropped capacity and viscosity with decreasing Δn for fixed $\frac{\Delta t}{\Delta n}$; (b) convergence of dropped capacity and viscosity by increasing $\frac{\Delta t}{\Delta n}$ for fixed Δn

3.2. Numerical results

Here, the same scenario considered in Section 2 is presented, but different discretizations, Δn and $\frac{\Delta t}{\Delta n}$ are considered. The simulation time is set to be 1350 [sec], which is slightly longer than t_s in Fig. 4(c) to ensure the convergence of the discharge flow rate; the number of simulated vehicles is 450. First, the dropped capacity, $\tilde{C}^-$, is calculated from numerical simulations of car-following models with given Δt and Δn. Table 1 summarizes the results with the dropped flow having units of veh/h. It is verified that different discretizations lead to different capacity drop ratios.

From (12), we find the viscosity present in each simulation, α, which is a function of Δn and $\frac{\Delta t}{\Delta n}$, and can be written as

$$\alpha = \left(1 - \frac{\gamma}{\tilde{C}^- \Delta\tau}\right) \frac{L}{\kappa \Delta\tau}, \tag{13}$$

where α, $\tilde{C}^-$ and γ depend on the discretization Δn and $\frac{\Delta t}{\Delta n}$. In particular, after obtaining the dropped capacity $\tilde{C}^-$ from simulations with given Δn and $\frac{\Delta t}{\Delta n}$, we first calculate γ from (12) and then α from the above equation. The corresponding values of the numerical viscosity are presented in parentheses in Table 1 in units of km^2/h. Fig. 5 clearly shows that the viscosity is directly related to the dropped capacity, and both increase with decreasing $\frac{\Delta t}{\Delta n}$ or increasing Δn. Furthermore, γ is shown to hardly depend on the discretization ($\gamma \sim 0.22$) in this numerical example. If one would assume it to be constant, the viscosity would have an hyperbolic relation with the dropped capacity, (13). Then, the theoretical dropped capacity can be approximated as $C^- = \frac{\gamma}{\Delta\tau}$, considering $\alpha = 0$. Further research is needed to have a better understanding of $\gamma(\Delta n, \frac{\Delta t}{\Delta n})$ and its relation to $\tilde{C}^-$.

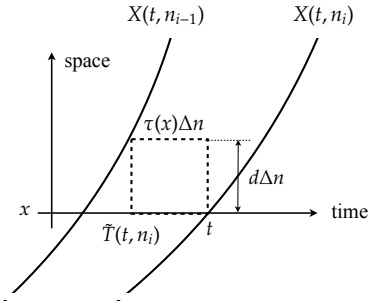


Fig. 6: Leader-follower relation in car-following mode ("Newell box")

We finally verify the assumption that $\tilde{\alpha} \ll 1$, since this dimensionless term varies between 0.0002 and 0.0200 in the cases analyzed. In conclusion, the numerical simulations verify analytic results for the viscosity. Moreover, the dropped capacity converges to the theoretical value based on the KW model of Jin (2018), i.e., $C^- = 1324.7$ [veh/h] for a fixed $\frac{\Delta t}{\Delta n}$ ratio and decreasing Δn.

4. Formation of capacity drop

The following sections clarify the properties of the formation of the capacity drop by examining a simple lead-vehicle problem where the initial vehicle drives at the free-flow speed and the following vehicles are initially equally spaced. Specifically, this section first shows that the spatial pattern of the traffic phase during the formation is the same as that for the stationary states (i.e., the KW model), i.e., the traffic state within and upstream of the bottleneck is in equilibrium while that downstream of the bottleneck is not in equilibrium when there is an upstream queue. We then describe the mechanism of the capacity drop. Furthermore, we establish a one-dimensional iterated function system model for the formation of the capacity drop. We prove a set of properties of a fixed point of the system (i.e., existence, uniqueness, stability, and speed of convergence).

Note that we here consider the continuum model (i.e., $\Delta t \to 0$ in Eq. (5) for a fixed $\frac{\Delta t}{\Delta n}$) because it is analytically tractable. We also assume that the vertical gradient $\Phi(x)$ is non-decreasing at $0 \le x \le L + d$[8], following Jin (2018).

4.1. Simple lead-vehicle problem

Let $n_i \equiv i\Delta n$ $(i = 0, 1, 2, \dots)$ denote the i-th vehicle[9]. The leader n_0 has the initial position is $X(0, n_0) = 0$, and drives with the free-flow speed u for time $t > 0$. The followers n_i $(i = 1, 2, \dots)$ are initially located with a spacing $s(0, n_i) = 1/k_0$ $(0 \le k_0 \le \kappa)$, and the initial position of vehicle n_i thus is $X(0, n_i) = -i\Delta n/k_0$. The followers then drive according to the model (5).

Let us first consider an almost obvious uncongested situation where $q_0 = Q(0, k_0) < C_2$. In this case, the followers have enough spacing and can pass through the bottleneck at the free-flow speed. The realized traffic flow rate for $x \ge 0$ is $q(t, x) = uk_0 = q_0$. Meanwhile, when $C_2 < q_0 = Q(0, k_0) \le C_1$, the follower n_i $(i = 1, 2, \dots)$ decelerates at some location $x < L$ because it starts to drive in the car-following mode[10]. Although it is not obvious which driving mode (i.e., free-flow, car-following, or bounded acceleration) will be realized downstream of the deceleration point, we have the following four lemmas if we assume that the followers are in the (congested) car-following mode inside the bottleneck (like in Fig. 4).

Lemma 4.1. *If traffic is congested inside the bottleneck, the car-following model (5) can be written as inhomogeneous Newell (2002)'s simplified car-following model with a location-dependent time gap:*

$$X(t, n_i) = X(\tilde{T}(t, n_i), n_{i-1}) - d\Delta n, \qquad \text{where} \quad \tilde{T}(t, n_i) = t - \tau(X(t, n_i))\Delta n. \tag{14}$$

[8] $\Phi(x)$ is assumed constant between $[L, L + d\Delta n]$, since d is usually very short.

[9] Δn here should be interpreted as a sampling interval from the continuum model that is different from a discretization unit for the discrete model.

[10] We use the words "mode" and "phase" interchangeably.

See Fig. 6 for an illustration.

Proof. We can obtain (14) by setting $\Delta t = \tau(X(t, n_i))\Delta n$ in (5), considering the congested car-following mode. ■

Lemma 4.2. *Assume that the vehicle n_i drives through the bottleneck $0 \le X(t, n_i) \le L$ in the car-following mode. Its equilibrium speed $v^E(t, n_i)$ is then given by*

$$v^E(t, n_i) = \left[\frac{1}{v(\tilde{t}, n_{i-1})} + \frac{\Delta n \Delta \tau}{L} \right]^{-1} \tag{15}$$

where $\tilde{t} \equiv \tilde{T}(t, n_i)$. This is the Del Castillo model for an inhomogeneous road (Del Castillo, 1996).

Proof. By taking the derivative of the expression in Eq.(14) with respect to t, we have

$$v^E(t, n_i) = v(\tilde{t}, n_{i-1}) \frac{\partial \tilde{T}(t, n_i)}{\partial t}, \quad \text{where} \quad \frac{\partial \tilde{T}(t, n_i)}{\partial t} = 1 - \frac{\Delta \tau \Delta n}{L} v^E(t, n_i). \tag{16}$$

After a simple manipulation, we obtain Eq.(15). ■

We here use the superscript "E" to indicate that the variables take equilibrium values.

Lemma 4.3. *Assume that the vehicle n_i drives through the bottleneck $0 \le X(t, n_i) \le L$ in the car-following mode. The equilibrium speed and acceleration of the vehicle n_i at time t are then equal to or less than the speed and acceleration, which are not necessarily in equilibrium, of the vehicle ahead n_{i-1} at time $\tilde{t}$, i.e.,*

$$v^E(t, n_i) \le v(\tilde{t}, n_{i-1}) \tag{17a}$$

$$a^E(t, n_i) \le a(\tilde{t}, n_{i-1}). \tag{17b}$$

For $v(\tilde{t}, n_{i-1}), a(\tilde{t}, n_{i-1}) > 0$, the equal sign holds only when $\Delta n \to 0$.

Proof. We first consider the reciprocal of the relation (15):

$$p^E(t, n_i) = p(\tilde{t}, n_{i-1}) + \frac{\Delta n \Delta \tau}{L}, \tag{18}$$

where $p(t, n_i) \equiv \frac{1}{v(t, n_i)}$ is the "pace." This equation means that the pace of the vehicle n_i at time t is that of the vehicle n_{i-1} at time $\tilde{t}$ plus a positive constant. Eq.(17a) thus holds true.

By taking the derivatives of both sides of Eq.(15) with respect to t, we have

$$a^E(t, n_i) = \frac{(v^E(t, n_i))^2 \cdot a(\tilde{t}, n_{i-1}) \cdot \frac{\partial \tilde{T}(t, n_i)}{\partial t}}{(v(\tilde{t}, n_{i-1}))^2} \quad \Leftrightarrow \quad \frac{a^E(t, n_i)}{a(\tilde{t}, n_{i-1})} = \frac{(v^E(t, n_i))^3}{(v(\tilde{t}, n_{i-1}))^3}, \tag{19}$$

Combining this result with Eq.(17a), we have Eq.(17b). ■

Eq.(17) can be interpreted as the "delay" of the acceleration due to the increase in the time gap because the equality holds when the time gap is constant.

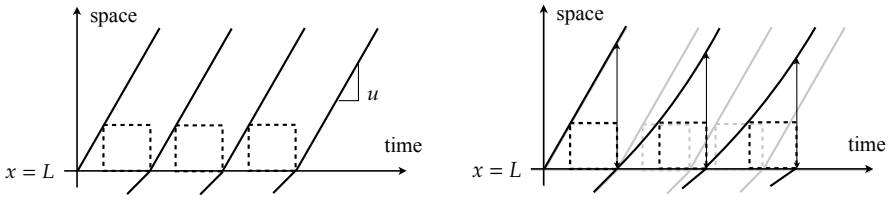
Lemma 4.4. *Assume that the vehicle n_i drives through the bottleneck $0 \le X(t, n_i) \le L$ in the car-following mode. The bounded acceleration constraint for the vehicle n_i at time t is then equal to or larger than that of the vehicle n_{i-1} at time $\tilde{t}$, i.e.,*

$$A(X(t, n_i), v^E(t, n_i)) \ge A(X(\tilde{t}, n_{i-1}), v(\tilde{t}, n_{i-1})) \tag{20}$$

Proof. For the constant bounded acceleration model, we have

$$a_0 - g\Phi(X(t, n_i)) \ge a_0 - g\Phi(X(\tilde{t}, n_{i-1}))$$

because $\Phi(X(t, n_i)) \le \Phi(X(\tilde{t}, n_{i-1}))$. By combining this with the property $\frac{\partial A(x,v)}{\partial v} \le 0$, the other natural bounded acceleration models (e.g., the TWOPAS model) satisfy Eq.(20) because of Eq.(17a). ■



(a) Newell's CF model (LWR model) (b) Proposed CF model

Fig. 7: Illustration of the formation of the capacity drop

We get the following theorem from the set of lemmas above.

Theorem 4.5. *Consider the simple lead-vehicle problem. Followers* n_i $(i = 1, 2, \dots)$ *then pass through the bottleneck* $0 \le X(t, n_i) \le L$ *in the equilibrium driving mode.*

Proof. When $q_0 \le C_2$, the followers apparently drive at the free-flow speed, i.e., the driving mode of the followers is in equilibrium. When $C_2 < q_0 \le C_1$, the followers have to decelerate at some location $x < L$. We let t_i be the time when the vehicle n_i decelerates. We here suppose that the vehicle n_i drives in the car-following mode inside the bottleneck after t_i. Then, from Eqs. (17b) and (20), we have the following equation no matter whether vehicle n_{i-1} is in the car-following or bounded acceleration mode[11].

$$a^E(t, n_i) \le a(\tilde{t}, n_{i-1}) \le A(X(\tilde{t}, n_{i-1}), v(\tilde{t}, n_{i-1})) \le A(X(t, n_i), v^E(t, n_i)) \qquad X(t_i, n_i) \le X(t, n_i) \le L,$$

This equation shows that the "equilibrium" acceleration rate, $a^E(t, n_i)$, never exceeds the bounded acceleration constraint, $A(X(t, n_i), v^E(t, n_i))$, inside the bottleneck. We thus conclude that the followers indeed drive in the car-following mode. ■

If the follower n_i continues to drive in the car-following mode $X(t', n_i) > L$, its speed recovers to the that of the vehicle ahead, i.e., $v^E(t', n_i) = v(\tilde{t}', n_{i-1})$. This is because the time gap $\tau(x)$ does not change $x > L$ and $\frac{\partial \tilde{T}(t, n_i)}{\partial t} = 1$ in the Eq.(16). Meanwhile, we know $v^E(t, n_i) < v(\tilde{t}, n_{i-1})$ at $X(t, n_i) = L$ (i.e., Eq.(17a) with $\Delta n > 0$). These discussions imply that, if the vehicle drives in the car-following mode, the instantaneous speed recovery (with an infinite acceleration rate) occurs at $X(t', n_i) > L$ unless the difference between $v^E(t', n_i)$ and $v(\tilde{t}', n_{i-1})$ becomes zero. However, this never happens in the proposed car-following model because of the bounded acceleration constraint. We thus have the following corollary.

Corollary 4.6. *Consider the simple lead-vehicle problem with* $C_2 < q_0 \le C_1$. *Then, for* $X(t, n_i) > L$, *the followers* n_i $(i = 1, 2, \dots)$ *drive in the bounded acceleration mode until the speed recovers to the free-flow speed.*

Theorem 4.5 and Corollary 4.6 show that the spatial pattern of the traffic mode for the the stationary states holds even during the transition states, like in Fig.4 (a). This property is useful for understanding the mechanism of the formation of the capacity drop and for constructing a simple model that describes the mechanism.

4.2. Mechanism of capacity drop

Fig. 7 illustrates the difference in the evolution of the headway between Newell (2002)'s car-following model and the proposed car-following model for the simple lead-vehicle problem with $C_2 < q_0 \le C_1$. In both models, the leader n_0 drives at the free-flow speed and the first follower n_1 enters the car-following mode at some location $x < L$. The speed of the first follower at position $x = L$ in both models is determined by the fundamental diagram (Theorem 4.5). The equilibrium speed is lower than the free-flow speed because the time gap increases inside the bottleneck (i.e., Eq.(17a) with $\Delta n > 0$).

[11] It is noted that Lemmas 4.1–4.4 do not impose any restriction on the driving mode of the vehicle ahead, n_{i-1}.

Let us first explain what happens in Newell (2002)'s car-following model, see Fig. 7(a). In this model, the speed of the first follower recovers to the free-flow speed instantaneously. The following vehicles n_i ($i = 2, 3, \dots$) then drive in the exactly the same way as the first follower, leading to a constant headway, which is the reciprocal of C_2.

Meanwhile, in the proposed car-following model, Fig. 7(b),the speed of the first follower cannot recover to the free-flow speed instantaneously owing to the bounded acceleration constraint (Corollary 4.6). The headway of the second follower at $x = L$ is therefore larger than that of Newell's car-following model, which implies that the flow rate becomes less than C_2. Additionally, if the speed of the first follower at $x = L + d\Delta n$ is lower than the free-flow speed, then the second follower's equilibrium speed at $x = L$ further decreases compared with that of the first follower. The headway of the third follower $x = L$ then further increases. This vicious cycle continues and the flow rate decreases (i.e., the headway increases). In other words, there are two types of "acceleration delays" due to the bounded acceleration and the increase in the time gap, and their feedback causes the capacity drop.

4.3. Iterated function system model for capacity drop

We construct an iterated function system model to reveal the properties of the formation of the capacity drop. Let v_i and $\tilde{v}_i$ respectively denote the speeds of the vehicles n_i ($i = 1, 2, \dots$) at $x = L$ and $x = L + d\Delta n$. On the basis of the mechanism mentioned above, we can predict the speed of vehicle n_{i+1} in the following two steps.

- Step 1. Bounded acceleration model: $\tilde{v}_i = G(v_i)$
- Step 2. Equilibrium relation: $v_{i+1} = F(\tilde{v}_i)$.

The function $F(\cdot)$ is no other than that given by Eq.(15). The function $G(\cdot)$ may not be obtained explicitly in general, but it is given for the constant bounded acceleration model, $A(x) = a_0 - g\Phi(x)$, as follows

$$\tilde{v}_i = G(v_i) \equiv \begin{cases} \sqrt{v_i^2 + 2[a_0 - g\Phi(L)] \cdot (d\Delta n)} & \text{if} \quad v_i < \overline{v} \\ u & \text{if} \quad v_i \geq \overline{v} \end{cases}, \tag{21}$$

where $\overline{v} \equiv \sqrt{u^2 - 2[a_0 - g\Phi(L)] \cdot (d\Delta n)}$. The case $v_i \geq \overline{v}$ may occur (especially, for the first follower) when the equilibrium speed v_i at $x = L$ is sufficiently close to the free-flow speed. However, this case is not interesting enough because the capacity drop does not evolve further. Assuming the constant bounded acceleration model, we therefore only address the case $v_1 = F(u) < \overline{v}$ in the following.

By employing the constant bounded acceleration model, we obtain an iterated function system (IFS) model for the capacity drop as the composition of the functions, $F(\cdot)$ and $G(\cdot)$.

$$v_{i+1} = f(v_i) \equiv F(G(v_i)) = \left(\alpha \Delta n + \frac{1}{\sqrt{\beta \Delta n + v_i^2}} \right)^{-1} \qquad i = 1, 2, \dots \tag{22a}$$

$$\text{where} \quad v_1 = F(u) < \overline{v}, \quad \alpha \equiv \frac{\Delta \tau}{L}, \quad \beta \equiv 2[a_0 - g\Phi(L)]d. \tag{22b}$$

Note that α and β are different from those in Section 3. The flow rate at the location $X(t, n_i) = L$ at time t can be calculated from the fundamental diagram: $q(t, L) = \frac{v_i}{s(t, n_i)}$.

Fig. 8 shows the (normalized) flow rate predicted using the IFS model. The parameter settings and Δn are the same as those for Fig. 4. Moreover, Fig. 8 reveals that the numerical plot (blue line) obtained using the proposed car-following model and the analytical prediction (yellow line) obtained using the IFS model match almost perfectly.

4.4. Properties of capacity drop formation

We first show the lemma regarding the mapping $f : [0, \overline{v}] \to [0, \overline{v}]$ of the IFS model.

Lemma 4.7. *The mapping $f : [0, \overline{v}] \to [0, \overline{v}]$ defined by Eq.*(22) *is a contraction mapping.*

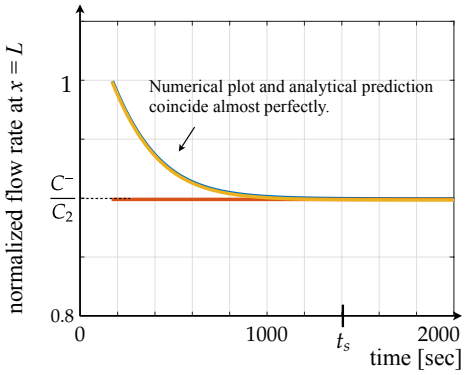


Fig. 8: Numerical example of IFS model (blue line: numerical plot; yellow line: analytical prediction ; orange line: normalized theoretical dropped capacity)

Proof. From the mean value theorem, for v_a and v_b ($v_a < v_b$), there is a speed $v_c \in (v_a, v_b)$ such that

$$\frac{f(v_b) - f(v_a)}{v_b - v_a} = f'(v_c) = \frac{v_c}{\left(\sqrt{\beta\Delta n + v_c^2}\right)^3 \left(\alpha\Delta n + \frac{1}{\sqrt{\beta\Delta n + v_c^2}}\right)^2}. \tag{23}$$

Also, for any $\alpha, \beta, \Delta n > 0$, it can be shown that the expression in Eq.(23) takes a value less than one. We thus have

$$\left|\frac{f(v_b) - f(v_a)}{v_b - v_a}\right| = |f'(v_c)| < 1,$$

which indicates that f is a contraction mapping. ■

From **Lemma 4.7**, we obtain the theorem regarding the following fixed point of the IFS model.

$$v^* = f(v^*) \in [0, \overline{v}] \tag{24}$$

Theorem 4.8. *The IFS model defined by Eq.*(22) *has a unique fixed point, and the fixed point is globally stable in the sense that the system* (22) *monotonically converges to it from any speed* $v \in [0, \overline{v}]$.

Proof. See, for examples, Scheinerman (1996) and Berinde (2007). ■

This theorem means that the proposed car-following model captures the important observation in the real traffic that the traffic states associated with the capacity drop are stable (Ozaki, 2003). When $\Delta n \to 0$, we further obtain the result that verifies the consistency between the proposed car-following model and the KW model (Jin, 2018).

Theorem 4.9. *When* $\Delta n \to 0$*, the traffic flow rate corresponding to the fixed point speed* (24) *coincides with the dropped capacity* C^- *of the kinematic wave model (Jin, 2018).*

Proof. From a simple manipulation, it is seen that the fixed point (24) is a non-negative real root of the following polynomial function:

$$\alpha^2\Delta n v^{*4} - 2\alpha v^{*3} + \alpha^2\beta(\Delta n)^2 v^{*2} - 2\alpha\beta\Delta n v^{*2} + \beta$$

Taking a limit $\Delta n \to 0$, we have $-2\alpha v^{*3} + \beta$. The fixed point speed is thus given by

$$v^* = \left(\frac{\beta}{2\alpha}\right)^{1/3} = \left(\frac{[a_0 - g\Phi(L)]dL}{\Delta\tau}\right)^{1/3} = \frac{y}{\kappa}. \tag{25}$$

Using the speed-spacing fundamental diagram, we finally obtain the corresponding traffic flow rate:

$$q^* = \frac{v^*}{s^*} = \frac{1}{\tau_2 + \frac{d}{v^*}} = \frac{y}{1 + y\tau_2} = C^-,$$

which is equivalent to Eq.(32) in Jin (2018). ■

We next investigate the speed of convergence of the IFS model. According to the contraction mapping principle (Berinde, 2007), we have the *a priori* error estimate:

$$|v_i - v^*| \leq \frac{M^i}{1 - M} \cdot |v_2 - v_1|, \tag{26}$$

where $M \equiv \max_{v\in[0,\bar{v}]} .|f'(v)| < 1$. This equation indicates that the rate of convergence of the system (22) is *exponential* and the upper bound of the number of iterations required to get the solution with an accuracy $|v_i - v^*| \leq \epsilon$ is given as

$$i \geq \frac{\ln \epsilon(1 - M) - \ln |v_2 - v_1|}{\ln M}. \tag{27}$$

From these equations, the speed of convergence increases with decreasing M and $|v_2 - v_1|$. Furthermore, we may say that the speed of convergence is dominated by $|v_2 - v_1|$ in early iterations and dominated by M in later ones.

Although it is difficult to obtain the values of the expressions in Eqs. (26) and (27) analytically, we can evaluate the effects of the system parameters α and β on $|v_2 - v_1|$ and M. For $\alpha < \alpha'$ and $\beta < \beta'$, we first obtain

$$M(\alpha, \beta) \geq M(\alpha', \beta), \quad M(\alpha, \beta) \geq M(\alpha, \beta'), \tag{28}$$

because $\frac{\partial f'(v;\alpha,\beta)}{\partial \alpha} < 0$ and $\frac{\partial f'(v;\alpha,\beta)}{\partial \beta} < 0$ (which can be easily confirmed from the functional form of Eq.(23)). Similarly, regarding $|v_2 - v_1| = F(u) - f(F(u))$, we have

$$\frac{\partial |v_2 - v_1|}{\partial \beta} = -\frac{\partial f(F(u); \alpha, \beta)}{\partial \beta} < 0, \tag{29}$$

but there is no monotone relation between $|v_2 - v_1|$ and α.

From Eqs. (28) and (29), we first conclude that the speed of convergence to the solution with the same accuracy increases with an increase in β (i.e., the increase in the bounded acceleration rate $[a_0 - g\Phi(L)]$ or the jam spacing d). Although the effect of α in early iterations could vary, the speed of convergence to a solution with a reasonable accuracy may increase with an increase in α (i.e., an increase in $\Delta\tau$ or a decrease in the length of the bottleneck L). It is interesting to highlight, that precisely these relations also hold to the capacity drop ratio, ζ, which increases with increasing $\Delta\tau$ and decreasing L.

Fig. 9 shows the speed of convergence of the system (22) to the solution with the accuracy $|v_i - v^*| < 10^{-2}$ [m/s] for different parameter settings. The base case setting is the same as the setting for Fig. 4, and the other four cases with $\tau_2 = 2.3, 2.5$ and $a_0 = 0.35, 0.40$ are calculated. We stop the calculation if the accuracy criterion is satisfied and use the value given by Eq.(25) (i.e., the fixed point speed for $\Delta n \to 0$) as v^*. Fig. 9 (a) shows the evolution of the normalized speed, v_i/u. The horizontal axis is the vehicle number i (or the number of iterations). This figure confirms that the conclusion of the above analysis is valid, i.e., the speed of convergence to the solution with the same accuracy increases with the increase in α (τ_2 in the present case) or β (a_0 in the present case). An interesting point to note about this figure is that, with an increase in α, the convergence speed increases even though the fixed point speed decreases. The right figure in Fig. 9 shows the evolution of the normalized queue discharge flow rate. The horizontal axis is time. This figure has the same tendency as the left figure, which implies that the above prediction may be applied to a more meaningful convergence process, the "time" evolution of the flow rate.

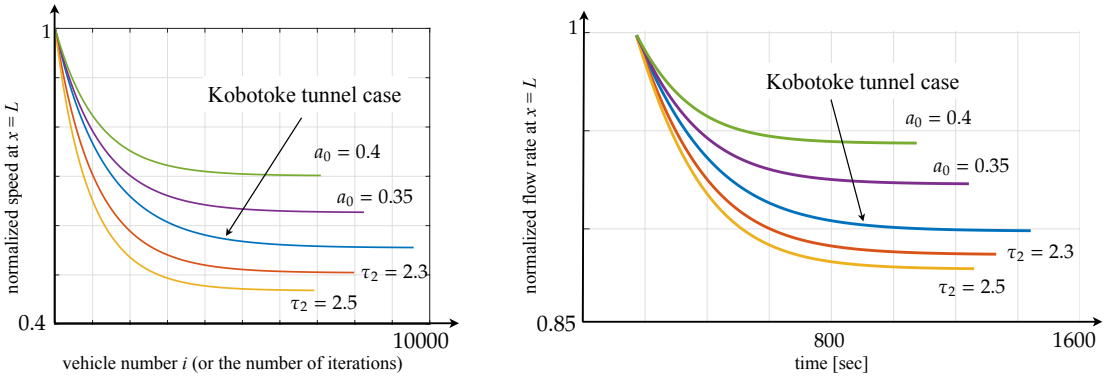


Fig. 9: Speed of convergence: (a) vehicle number vs. normalized speed; (b): time vs. normalized queue discharge flow rate

5. Impacts of heterogeneous drivers on capacity drop

We have so far analyzed the formation and stabilization process of the capacity drop, which was not possible using the KW model. Another advantage of the proposed car-following model (and the iterated function system) over the KW model is that heterogeneous drivers can be easily incorporated. This section thus analyzes the effect of heterogeneous drivers on the capacity drop. This refinement is valuable to check the robustness (Section 5.1) of our conclusions in the homogeneous analysis. Furthermore, the heterogeneous model is more realistic and less rigid, and makes a new theoretical prediction that can be checked by numerical experiments (Section 5.2).

To introduce the heterogeneity of drivers, we let some parameters of the model be random variables. We write the random parameters as functions of ω, where $\Omega \ni \omega$ is the random space. More specifically, the random parameters $\{\tau_{2,i}(\omega), d_i(\omega), L_i(\omega), a_{0,i}(\omega)\}$ of the car-following model and the resulting parameters $\{\alpha_i(\omega), \beta_i(\omega)\}$ are assumed to be independent and identically distributed. The supports of the parameters distributions are bounded from both above and below by positive and non-infinite constants to ensure that the parameters values are physically meaningful.

5.1. Heterogeneous speed-spacing relations

We consider simple lead-vehicle problems where there is a sufficient traffic demand so that a queue always exists after the first following vehicle n_1 passing through the bottleneck. In this subsection, we assume that only $\{\tau_{2,i}(\omega), d_i(\omega)\}$ are random (i.e., the speed-spacing relation is heterogeneous). Thus, $\alpha_i(\omega) = (\tau_{2,i}(\omega) - \tau_1)/L$ and $\beta_i(\omega) = 2[a_0 - g\Phi(L)]d_i(\omega)$ are also random. In such a situation, Theorem 4.5 and Corollary 4.6 still hold. We can therefore analyze the stochastic formation of the capacity drop through the following iterated random function system (IRFS) model.

$$v_{i+1} = f(v_i; \tilde{\omega}) = \left(\alpha_{i+1}(\omega')\Delta n + \frac{1}{\sqrt{\beta_i(\omega)\Delta n + v_i^2}} \right)^{-1} \qquad i = 1, 2, \ldots \tag{30}$$

where $\tilde{\omega} \equiv (\omega, \omega') \in \Omega \times \Omega,\ v_1 = F(u; \omega) < \min\{\overline{v}(\omega),\ \omega \in \Omega\}$

This model clearly has the Markov property that the conditional distribution of v_{i+1} given $\tilde{\omega}$ depends only on v_i. The Markov chain induced by this model is analyzed in the following.

Because $f(\cdot, \tilde{\omega})$ for each $\tilde{\omega}$ is a contraction mapping as in the deterministic system model (22), the IRFS model (30) monotonically converges to $v^*(\tilde{\omega})$ from $[0, v_1]$ for a fixed $\tilde{\omega}$ (we will refer to this property as "Property A"). There also exists some $M(\tilde{\omega})$ for $v_a, v_b \in [0, v_1],\ v_a \neq v_b$ such that

$$\left| \frac{f(v_b, \tilde{\omega}) - f(v_a, \tilde{\omega})}{v_b - v_a} \right| < M(\tilde{\omega}) < 1.$$

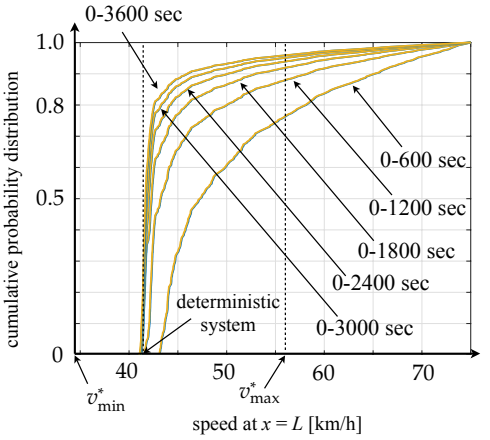


Fig. 10: Comparison of cumulative probability distributions of real vehicles' speeds for case of heterogeneous speed-spacing relations (blue line: proposed car-following model; yellow line: IRFS model)

Therefore, $f(\cdot, \tilde{\omega})$ satisfies the condition $\mathbb{E}[\log M(\tilde{\omega})] < 0$, the so-called "*contraction on average*" (Diaconis and Freedman, 1999). We then obtain the following theorem.

Theorem 5.1. *The Markov chain induced by the IRFS* (30) *has a unique stationary distribution.*

Proof. According to Theorem 1.1 stated by Diaconis and Freedman (1999), for there to be a unique stationary distribution of the induced Markov chain, $f(\cdot, \tilde{\omega})$ needs to satisfy the following regularity conditions as well as be contracting on average:

$$\mathbb{E}[M(\tilde{\omega})] < \infty, \quad \mathbb{E}[|f(v, \tilde{\omega}) - v|] < \infty \text{ for some } v \in [0, v_1].$$

The function $f(\cdot, \tilde{\omega})$ obviously satisfies these conditions. ∎

While Theorem 5.1 does not provide the probability distribution in the interval $[0, v_1]$ when $i \to \infty$, we can extract such information using the uniqueness property and Property A. Let us consider the case that the initial speed is v_1. Then, from Property A, it is seen that (i) the speed decreases monotonically until it reaches $v^*_{\max} \equiv \max(v^*(\tilde{\omega}))$, and (ii) the speed does not deviate from $[v^*_{\min}, v^*_{\max}]$ after that. Here, $v^*_{\min} \equiv \min(v^*(\tilde{\omega}))$. We therefore expect that the probability concentrates in the interval $[v^*_{\min}, v^*_{\max}]$ when $i \to \infty$. This expectation should not depend on the initial speed because the stationary distribution is unique. If the minimum values of $\alpha(\omega)$ and $\beta(\omega)$ are 80% of the maximum values and $v^*_{\min} = 30$ [km/h], we can calculate the length of the interval as

$$\frac{v^*_{\max} - v^*_{\min}}{v^*_{\min}} = \left(\frac{\alpha_{\max}\beta_{\max}}{\alpha_{\min}\beta_{\min}}\right)^{1/3} - 1 = \left(\frac{1}{0.64}\right)^{1/3} - 1 \approx 0.16 \quad \Rightarrow \quad |v^*_{\max} - v^*_{\min}| \approx 4.8 \text{ [km/h]}$$

where we use the fixed point speed (25).

We present the numerical results of the IRFS model (30). The random parameters $\{\tau_{2,i}(\omega), d_i(\omega)\}$ are assumed to follow beta distributions whose supports are $[0.85\mathbb{E}[\tau_{2,i}(\omega)], 1.15\mathbb{E}[\tau_{2,i}(\omega)]]$ and $[0.85\mathbb{E}[d_i(\omega)], 1.15\mathbb{E}[d_i(\omega)]]$[12]; the two positive shape parameters of the distribution are set as 2. The mean values of the random parameters and other parameters are the same as those in Fig. 4. (i.e., $\mathbb{E}[\tau_{2,i}(\omega)] = 2.1$ [sec] and $\mathbb{E}[d_i(\omega)] = 1/140$ [km/veh]). For these settings, $v^*_{\max} \approx 56.3$ [km/h] and $v^*_{\min} \approx 34.5$ [km/h]. Fig. 10 shows the cumulative probability distributions of speed at $x = L$ for different time intervals obtained using the proposed car-following model (blue line) and the IRFS model (yellow line). As in the deterministic case (Fig. 8), we see that the IRFS model well describes the stochastic formation of the capacity drop obtained using the proposed car-following model. The figure also reveals

[12] In the model implementation, the parameters for real vehicle i with integer number n_i will be assigned to imaginary vehicles i' that satisfy $n_{i'} \in (n_i - 1, n_i)$.

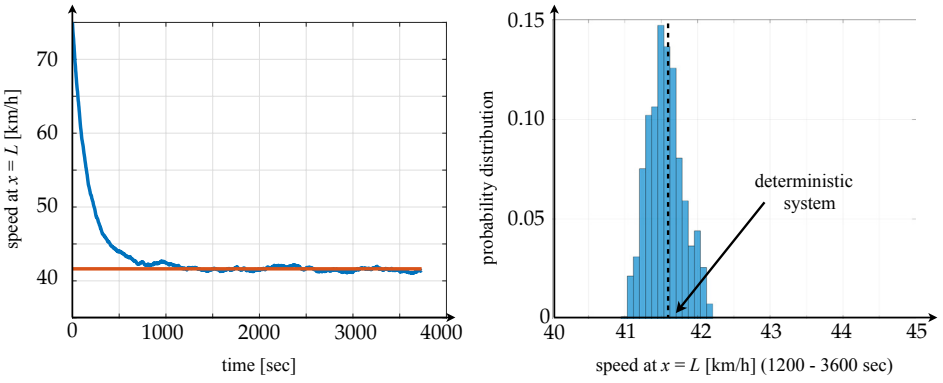


Fig. 11: Convergence process and stationary distribution for case of heterogeneous speed-spacing relations: (a) evolution of real vehicles' speed (red line: fixed point speed of the deterministic model); (b) probability distribution during 1200 − 3600 [sec])

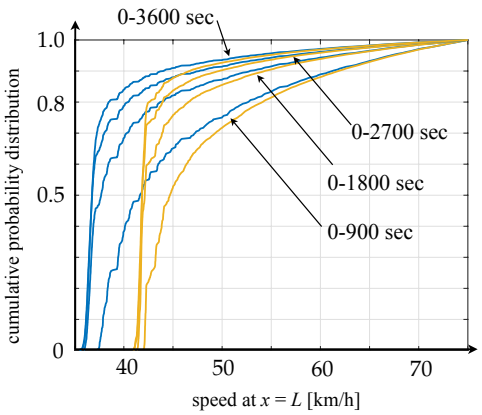


Fig. 12: Comparison of cumulative probability distributions of speed for case with heterogeneous acceleration bounds (blue line: proposed car-following model; yellow line: IRFS model)

that the probability concentrates within the somewhat wide interval $[v^*_{\min}, v^*_{\max}]$ as time proceeds. Fig. 11(a) shows the convergence process of the model (30). After around 1200 [sec], the speed remains close to the fixed point speed of the deterministic model (i.e., Eq.(22) with the mean parameters values). This transition time is almost the same with the deterministic case, t_s, and is comparable with the empirical observation. The corresponding probability (density) distribution is shown in Fig. 11(b).

The theoretical and numerical analyses suggest that the basic nature of the capacity drop presented in the previous section (i.e., uniqueness, stability and transition time) is inherited when the speed-spacing relations are heterogeneous. Conversely, with this heterogeneity only, i.e., $\{\tau_{2,i}(\omega), d_i(\omega)\}$ a strong effect on the capacity drop stationary state and its formation process is not expected.

5.2. Heterogeneous acceleration bounds and moving bottleneck phenomenon

We finally analyze the effects of the heterogeneous acceleration bounds on the capacity drop. This case is more complicated than the one in the previous subsection because Eq.(20) may be violated and, as a result, Theorem 4.5 and Corollary 4.6 cannot be expected to be hold in general[13]. We may need to analyze the proposed continuum car-following model directly instead of the IRFS model (30).

[13] If the parameter L is randomly distributed (i.e., individual recognition of the length of the bottleneck is heterogeneous), the spatial pattern of the traffic mode can also differ from that predicted by Theorem 4.5 and Corollary 4.6.

To illustrate this issue, we present a numerical example in which the parameter $\{a_{0,i}(\omega)\}$ is assumed to follow beta distribution whose support is $[0.85\mathbb{E}[a_{0,i}(\omega)], 1.15\mathbb{E}[a_{0,i}(\omega)]]$ and its shape parameters are set as 2, and other parameters of the car-following model are fixed. The proposed car-following model (blue line) and the IRFS model (yellow line) are compared in Fig. 12. While the two cumulative probability distributions matches well in the high-speed range (say, above 60 [km/h]), there are systematic differences in the low-speed range.

This phenomenon can be explained as follows. Let us consider the case that the vehicle n_{i-1} switches from the car-following mode to the bounded acceleration mode at the end of the bottleneck $x = L$, and accelerates at the rate, $a_{i-1,0}$, which exceeds the rate of the following vehicle, $a_{i,0}$ (i.e., Eq.(20) is violated). Then, from Eq.(19), the equilibrium acceleration rate of the vehicle n_i at $X(t, n_i) < L$ is given by

$$a^E(t, n_i) = \frac{(v^E(t, n_i))^3}{(v(\tilde{t}, n_{i-1}))^3} \left[a_{i-1,0} - g\Phi(X(\tilde{t}, n_{i-1}))\right],$$

where we assume $X(\tilde{t}, n_{i-1}) > L$. When the speed is sufficiently low, their speed difference can be small, i.e., $v^E(t, n_i) \approx v(\tilde{t}, n_{i-1})$. Thus, $a^E(t, n_i) \approx a_{i-1,0} - g\Phi(X(\tilde{t}, n_{i-1})) > a_{i,0} - g\Phi(X(t, n_i))$ happens, which means that the vehicle n_i is in the BA mode at $X(t, n_i) < L$ and the speed at $x = L$ becomes lower than the equilibrium speed. Furthermore, this slowly accelerating vehicle becomes a *moving bottleneck* and creates a "void" in its front and may systematically reduce the overall traffic speed and flow rate.

While this theoretical prediction should be tested by empirical experiments, it seems to be related to the following phenomenon observed at some sags and tunnels: even during the stabilized congestion period, the queue discharge flow rate keeps slightly decreasing as the time that the drivers have spent in the queue increases (Koshi, 2003). In the reference, this phenomenon was explained by a human psychology that *drivers tend to loose more their tension in the following cars ahead as they have been caught in the queue longer* (the effect may be represented in the model with a queue-length-dependent a_0). The psychological effect may not be expected for *all* drivers but the above prediction indicates that *some* drivers with this effect can become the moving bottlenecks and lead to the decrease in the queue discharge flow rate.

6. Conclusion

We presented a continuum car-following model of the capacity drop at sag and tunnel bottlenecks by extending the BA-LWR model (Jin and Laval, 2018) with inhomogeneous fundamental diagrams. In Figs. 3 and 4, we demonstrated that the stationary speed profiles, the low acceleration rates, the capacity drop ratio, and the degree of the transition time in the model are consistent with empirical observations at the Kobotoke tunnel in Japan.

We then investigated the impacts on the stationary states and dropped capacity of the numerical viscosity caused by the discretization method and showed that the dropped capacity converges to the theoretical value given by the KW model of Jin (2018). We further analyzed the formation of the capacity drop. Based on the spatial pattern of the traffic phases that arises in a lead-vehicle problem, we clearly described the mechanism of the capacity drop formation, i.e., there are two types of acceleration delays due to the bounded acceleration and the increase in the time gap, and their feedback causes the capacity drop. We modeled the feedback process as a one-dimensional iterated function system with a contraction mapping. With this desirable property, we demonstrated that (i) the capacity drop stationary state is globally stable and the rate of convergence is exponential; (ii) the stability and convergence property holds even when the speed-spacing relations are heterogeneous (i.e., the stationary distribution of the Markov chain induced by the model is unique); and (iii) heterogeneous acceleration bounds may generate moving bottlenecks and reduce the speed of the capacity drop stationary state.

The insights on the temporal on-rise of capacity drop inside the sag and tunnel bottleneck can be used to design better freeway control and vehicle-based guidance strategies to prevent, delay, or eliminate capacity drops at critical bottlenecks. As one example in this direction, Martínez and Jin (2018) analyzed the effect that the variable speed limit (VSL) application area has on the control effectiveness, and analytically obtained an optimal location of the speed limit application area.

In future work, we will be interested in extending the models and methods proposed in this study to analyze the more complicated phenomena observed at sags and tunnels (e.g., the probabilistic capacity reduction, stop-and-go traffic upstream of bottlenecks, and the queue-length dependent reduction of the queue discharge flow rate) and the

capacity drop phenomenon at other bottlenecks[14]. We will also be interested in calibrating and validating the models with empirical data, especially with respect to the temporal transition from a free-flow condition to the capacity drop stationary states.

Acknowledgments

The authors express their gratitude to four anonymous referees for their careful reading of the manuscript and useful suggestions. The first author would like to thank the Kajima Foundation's Assistance for Exchange of Researcher for the support. The second author would like to thank the Graduate Balsells Fellowship program for the support. The third author would like to thank NSF/CMMI #1434753 and UCTC for their support.

References

Allen, R., Harwood, D., Chrstos, J., Glauz, W., 2000. The Capability and Enhancement of VDANL and TWOPAS for Analyzing Vehicle Performance on Upgrades and Downgrades within IHSDM. Technical report. Federal Highway Administration 00-078.

Bando, M., Hasebe, K., Nakayama, A., Shibata, A., Sugiyama, Y., 1995. Dynamical model of traffic congestion and numerical simulation. Physical Review E 51, 1035–1042.

Banks, J.H., 1991. Two-capacity phenomenon at freeway bottlenecks: A basis for ramp metering. Transportation Research Record , 83–90.

Berinde, V., 2007. Iterative Approximation of Fixed Points. Springer.

Brilon, W., Bressler, A., 2004. Traffic flow on freeway upgrades. Transportation Research Record 1883, 112–121.

Cassidy, M.J., Bertini, R.L., 1999. Some traffic features at freeway bottlenecks. Transportation Research Part B 33, 25–42.

Daganzo, C.F., 1995. Requiem for second-order fluid approximations of traffic flow. Transportation Research Part B 29, 277–286.

Del Castillo, J.M., 1996. A car following model based on the Lighthill-Whitham theory, in: Lesort, J.B. (Ed.), Proceedings of the 13th International Symposium on Transportation and Traffic Theory, Pergamon. pp. 517–538.

Diaconis, P., Freedman, D., 1999. Iterated random functions. SIAM Review 41, 45–76.

Edie, L.C., Foote, R.S., 1958. Traffic flow in tunnels, in: Highway Research Board Proceedings, pp. 334–344.

Furuichi, T., Yamamoto, S., Kotani, M., Iwasaki, M., 2003. Characteristics of spatial speed change at motorway sag, in: Transportation Research Board 2003 Annual Meeting, pp. 1–23.

Goñi Ros, B., Knoop, V.L., van Arem, B., Hoogendoorn, S.P., 2013. Car-following behavior at sags and its impacts on traffic flow, in: The 92nd TRB Annual Meeting.

Goñi Ros, B., Knoop, V.L., van Arem, B., Hoogendoorn, S.P., 2014. Empirical analysis of the causes of stop-and-go waves at sags. IET Intelligent Transport Systems 8, 499–506.

Goñi Ros, B., Knoop, V.L., Shiomi, Y., Takahashi, T., van Arem, B., Hoogendoorn, S.P., 2016. Modeling traffic at sags. International Journal of Intelligent Transportation Systems Research 14, 64–74.

Hall, F.L., Agyemang-Duah, K., 1991. Freeway Capacity Drop and the Definition of Capacity. Transportation Research Record , 91–98.

Hatakenaka, H., Hirasawa, T., Yamada, K., Yamada, H., Katayama, Y., Maeda, M., 2006. Development of ahs for traffic congestion in sag sections, in: Proceedings of the 13th ITS World Congress.

Japan Highway Public Corporation, 1987. A study on traffic capacity at tunnels on the Chuo Expressway. Technical Report (in Japanese).

Jin, W.L., 2016. On the equivalence between continuum and car-following models of traffic flow. Transportation Research Part B 93, 543–559.

Jin, W.L., 2017a. A first-order behavioral model of capacity drop. Transportation Research Part B 105, 438–457.

Jin, W.L., 2017b. Kinematic wave models of lane-drop bottlenecks. Transportation Research Part B 105, 507–522.

Jin, W.l., 2017c. Nonstandard second-order formulation of the LWR model. URL: http://arxiv.org/abs/1701.08926.

Jin, W.L., 2018. Kinematic wave models of sag and tunnel bottlenecks. Transportation Research Part B 107, 41–56.

Jin, W.L., Laval, J., 2018. Bounded acceleration traffic flow models: A unified approach. Transportation Research Part B 111, 1–18.

Komada, K., Masukura, S., Nagatani, T., 2009. Effect of gravitational force upon traffic flow with gradients. Physica A 388, 2880–2894.

Koshi, M., 2003. An interpretation of a traffic engineer on vehicular traffic flow, in: Traffic and Granular Flow'01. Springer Berlin Heidelberg, pp. 199–210.

Koshi, M., Iwasaki, M., Ohkura, I., 1983. Some findings and an overview on vehicular flow characteristics, in: Hurdle, V.F., Hauer, E., Steuart, G.N. (Eds.), Proceedings of the 8th International Symposium on Transportation and Traffic Theory, pp. 403–426.

Koshi, M., Kuwahara, M., Akahane, H., 1992. Capacity of sags and tunnels on Japanese motorways. ITE Journal 62, 17–22.

Laval, J.A., Leclercq, L., 2008. Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model. Transportation Research Part B 42, 511–522.

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

[14] See, for example, Koshi et al. (1983); Banks (1991); Hall and Agyemang-Duah (1991); Cassidy and Bertini (1999); Jin (2017a,b) and references therein.

Leclercq, L., Laval, J.A., Chevallier, E., 2007. The Lagrangian coordinates and what it means for first order traffic flow models, in: Allsop, R., Bell, M., Heydecker, B. (Eds.), Proceedings of the 17th International Symposium on Transportation and Traffic Theory, Elsevier. pp. 735–753.
Martínez, I., Jin, W.L., 2018. Impact of VSL location on capacity drop: A case of sag and tunnel bottlenecks. Transportation Research Procedia 34, 12–19.
Moriyama, Y., Mitsuhashi, M., Hirai, S., Oguchi, T., 2011. The effect on lane utilization and traffic capacity of adding an auxiliary lane. Procedia - Social and Behavioral Sciences 16, 37–47.
Newell, G.F., 2002. A simplified car-following theory: a lower order model. Transportation Research Part B 36, 195–205.
Oguchi, T., 2004. Queued traffic flow occurrence at capacity bottlenecks on expressway basic sections, in: The 83rd TRB Annual Meeting.
Oguchi, T., Konuma, R., 2009. Comparative study of car-following models for describing breakdown phenomena at sags, in: Proceedings of the 16th World Congress on ITS.
Okamura, H., Watanabe, S., Izumi, M., 2000a. Traffic phenomena at head of the queue on basic intercity motorway sections. JSTE Traffic Engineering 35, 48–56 (in Japanese).
Okamura, H., Watanabe, S., Watanabe, T., 2000b. An empirical study on the capacity of bottlenecks on the basic suburban expressway sections in Japan, in: Transportation Research Circular E-C018, pp. 120–129.
Ozaki, H., 1994. Car-following behavior and bottleneck phenomena at vertical alignment sags. Doctor's dissertation, The University of Tokyo (in Japanese).
Ozaki, H., 2003. Modeling of vehicular behavior from road traffic engineering perspectives, in: Traffic and Granular Flow'01. Springer Berlin Heidelberg, pp. 281–292.
Patire, A.D., Cassidy, M.J., 2011. Lane changing patterns of bane and benefit: Observations of an uphill expressway. Transportation Research Part B 45, 656–666.
Scheinerman, E.R., 1996. Invitation to Dynamical Systems. Prentice-Hall.
Sun, J., Li, T., Yu, M., Zhang, H.M., 2018. Exploring the congestion pattern at long-queued tunnel sag and increasing the efficiency by control. IEEE Transactions on Intelligent Transportation Systems , 1–10.
Treiber, M., Hennecke, A., Helbing, D., 2000. Congested Traffic States in Empirical Observations and Microscopic Simulations. Physical Review E 62, 1805–1824.
Xing, J., Koshi, M., 1995. A study on the bottleneck phenomena and car-following behavior on sags of motorways. Journal of the Japan Society of Civil Engineers IV-5, 45–55 (in Japanese).
Xing, J., Muramatsu, E., Harayama, T., 2014. Balance lane use with VMS to mitigate motorway traffic congestion. International Journal of Intelligent Transportation Systems Research 12, 26–35.
Yoshizawa, R., Shiomi, Y., Uno, N., Iida, K., Yamaguchi, M., 2012. Analysis of car-following behavior on sag and curve sections at intercity expressways with driving simulator. International Journal of Intelligent Transportation Systems Research 10, 56–65.