

Behavior-rule dependence of capacity gains in autonomous/human mixed cellular-automaton traffic, and the absence of a penetration threshold for slow-to-start hysteresis suppression

Mitsuyoshi Yagyu^{a,*}

^aDepartment of Information and Data Science, Faculty of Information and Data Science, Notre Dame Seishin University, 2-16-9 Ifuku-cho, Kita-ku, Okayama City 700-8516 Japan

Abstract

The dependence of capacity gains on the choice of autonomous-vehicle (AV) behavior rule is investigated using a Nagel–Schreckenberg-type cellular-automaton model of mixed AV and human-driven vehicle (HV) traffic. Four AV following rules (a naive deterministic rule, the Intelligent Driver Model, adaptive cruise control, and cooperative adaptive cruise control) are implemented under a common collision-safety architecture, in which each vehicle computes a desired speed from its own rule and is then subject to the same hard safety floor. Slow-to-start dynamics are retained for the human-driven vehicles so that metastability and hysteresis are present, and the hysteresis loop width is measured as a function of AV penetration rate on a fine penetration grid at four ring sizes. The framework is validated against an exact two-species disordered exclusion-process solution at unit maximum velocity, with agreement to within 0.53 percent.

The capacity gain from AV penetration is found to be highly sensitive to the calibration of the AV following rule rather than being an intrinsic property of the rule itself: under human-calibrated Intelligent Driver Model parameters, capacity is not monotonic in penetration but falls below the human-only baseline at intermediate penetration and recovers above it at full penetration, so that a mixture of the two populations carries less flow

*Corresponding author.

Email address: mitsu.yagyu@gmail.com (Mitsuyoshi Yagyu)

than either population alone, whereas doubling the acceleration and deceleration capability removes the deficit. Cooperative adaptive cruise control, in contrast, saturates the common safety floor and becomes indistinguishable from the naive rule at high penetration, while adaptive cruise control remains distinct owing to its longer effective headway. The suppression of slow-to-start hysteresis, by contrast, is shown to be a continuous crossover rather than a threshold effect: the loop width decreases monotonically within uncertainty across the penetration range, the decline accelerates above roughly seventy percent penetration, and the curves for ring sizes spanning a factor of eight agree to within six hundredths of a loop width at every penetration and differ by only three percent in their mean, with the larger rings giving slightly smaller widths, so that the descent does not sharpen as the system is enlarged. This part of the study uses the naive deterministic AV rule alone, so it establishes the shape of the suppression and not its dependence on the behavior rule. No critical penetration rate is therefore identified. It is shown further, by direct measurement, that the AV share of standing vehicles tracks the penetration rate to within five parts in ten thousand and that standing times are indistinguishable between the two vehicle types, which excludes an explanation resting on AVs coming to dominate the standing population. We conclude that the widely assumed benefit of AV penetration for traffic flow is not a model-independent conclusion for capacity, where it depends sensitively on the behavior rule and its calibration, and that for hysteresis it accrues gradually rather than at a deployment threshold.

Keywords: cellular automaton, Nagel–Schreckenberg model, autonomous vehicles, mixed traffic, car-following models, hysteresis

Note on this version

This is the second version of this preprint. The first version was posted on 25 July 2026 as doi:10.31224/7696. Additional computation since then has changed two of the reported results and revealed four errors. The following corrections, together with a narrowing of the framing, have been made, and readers of version 1 should note them.

1. Version 1’s title claimed behavior-rule dependence for both the capacity gains and the hysteresis suppression. Only the capacity comparison was run across the four AV rules; the hysteresis scan used the naive

deterministic rule alone. The claim was therefore not supported by the work, and the title, the abstract, and the closing discussion have been changed to say so.

2. Version 1 reported that hysteresis suppression is “threshold-like rather than proportional”, with the loop width “largely unchanged up to $r \approx 0.4$ ” and an effect “concentrated above roughly 60% penetration”. A repeat of that scan on a penetration grid eight times finer, with twenty seeds instead of three and at four ring sizes instead of one, does not support this. The suppression is a continuous crossover with no critical penetration rate and no size-dependent sharpening, and the coexistence region narrows steadily from $r \approx 0.2$ onward rather than staying flat up to $r \approx 0.4$. The non-monotonicity of the original six-point table was run-to-run noise. Section 3.2 has been rewritten accordingly.
3. Version 1 explained the shape of that curve by asserting that AVs must come to dominate the population of standing vehicles. Direct measurement (Section 3.2.2) shows that the AV share of standing vehicles equals the penetration rate to within 5×10^{-4} at every r , and that mean standing times are indistinguishable between the two vehicle types. The assertion was incorrect and has been replaced by the measurement.
4. Version 1 stated the IDM acceleration in lattice units as $a \approx 0.13$ cells/step². The correct conversion is $0.73/7.5 \approx 0.097$ cells/step². The simulations always used the correct value; only the text was wrong.
5. In version 1 the continuous rules (IDM, ACC, CACC) were given the center-to-center lattice distance to the leader, whereas those models are formulated in net bumper-to-bumper spacing and the shared safety floor used the net spacing. The rules were therefore given one cell (7.5 m) more room than intended. This has been corrected and the axis-1 scan rerun under both conventions. The effect is substantial: capacity at full penetration changes by +53% for IDM and −22% for ACC. The reported capacities and the statement of Finding 1 change accordingly; see Sections 2.6 and 3.1.1.
6. Version 1’s abstract stated that two independent methods were used to quantify the loop width as a function of r . The adiabatic-ramp method was applied only to the $r = 0$ baseline, as Section 2.9 correctly described; the abstract has been corrected.
7. The framing of the contribution has been narrowed, and Jiang et al. (2021), Jiang and Ahn (2026) and Kinjo and Tomoeda (2025) are now cited and positioned in Section 1.

1. Introduction

The Nagel–Schreckenberg (NaSch) cellular automaton (CA) (Nagel and Schreckenberg, 1992) and its many variants (Chowdhury et al., 2000) remain a standard, computationally efficient tool for studying macroscopic traffic flow phenomena: the triangular flow-density fundamental diagram, jam formation, and, in slow-to-start variants, metastability and hysteresis. With the emergence of autonomous and connected vehicles, a large body of work has asked how the AV penetration rate r , equivalently the AV fraction, reshapes the fundamental diagram, typically by giving AVs a deterministic (non-stochastic) update rule, sometimes combined with a reduced following gap. This question, AV penetration rate versus fundamental diagram, is by now well covered by prior studies using a variety of specific AV control laws, including cyber-interaction-based communication rules (Guo et al., 2023), a fixed CACC control law combined with a calibrated human car-following model (Ma et al., 2023), and reinforcement-learning-trained AV policies (Poudel et al., 2024; Kumar et al., 2026). A simple re-derivation of the same qualitative conclusion, that AVs improve capacity, under yet another AV rule would add little.

Comparisons across control laws within a single CA have themselves been made before. Jiang et al. (2021) build one cellular automaton containing three car-following modes, human driving together with adaptive and cooperative adaptive cruise control, and report the fundamental diagram, congestion, and speed volatility across CAV penetration rates; Kinjo and Tomoeda (2025) compare gap-based and flow-based control strategies inside a single controlled stochastic CA, albeit not in mixed traffic. We do not claim, therefore, that comparing behavior rules on one lattice is itself new. What we add is a controlled comparison in which every rule is placed under one explicit, shared collision-safety architecture, so that differences between rules cannot be attributed to differences in how each rule was made safe, and in which the resulting architecture is separately verified against an exact solution. Within that setting we ask: how much does the *answer itself* depend on which AV behavior model is assumed? We implement four car-following rules on the same NaSch lattice, namely a simple deterministic rule (no randomness, otherwise identical to a human-driven vehicle, HV), the Intelligent Driver Model (IDM; Treiber et al., 2000), and the adaptive and cooperative adaptive cruise control laws (ACC and CACC) calibrated directly from experimental vehicle data by Milanés and Shladover (2014) (a calibration also

connected to macroscopic flow characteristics by van Arem et al. (2006)), under one common, collision-safe update architecture, and compare their effect on the fundamental diagram across AV penetration rate r .

A second, related but distinct question is whether AV penetration affects metastability and hysteresis. Plain (non-slow-to-start) NaSch traffic does not exhibit hysteresis; the effect requires a slow-to-start, velocity-dependent-randomization (VDR) rule in which standing vehicles restart with a different, typically lower, probability than moving vehicles brake (Barlović et al., 1998; Maerivoet and De Moor, 2004). The effect of AV penetration on traffic hysteresis has been studied in the sense of the oscillatory, dynamic fundamental diagram: Jiang and Ahn (2026) formulate a sequence-based stochastic dynamic fundamental diagram for mixed platoons and report both that AV/HV sequencing significantly alters the size of the hysteresis loop and that higher AV shares dampen its magnitude and variability. The quantity at issue here is a different one. Slow-to-start metastability is the coexistence of two branches, a metastable free-flow branch and a jammed branch, in the sense of Barlović et al. (1998), rather than a loop traced out by a platoon responding to a disturbance. How AV penetration closes *that* coexistence, and in particular whether it closes gradually or at a critical penetration rate, has to our knowledge not been reported. The distinction matters practically, because a critical penetration rate would imply that the benefit is unavailable until a deployment threshold is crossed, whereas a gradual closure implies that it accrues from the first AV onward.

Implementing multiple literature AV models under mixed traffic surfaced a methodological pitfall worth reporting on its own terms (Section 2.3): a common way of representing an AV’s advantage, letting it tolerate a shorter following gap than an HV, is only collision-free under NaSch’s fully parallel update rule if the vehicle ahead is guaranteed to move every timestep in which it might be adjacent to a follower. This holds for the standard single-cell safety margin but not for a reduced one, and the violation is invisible in a homogeneous AV population, where it happens to be safe, while occurring in a large fraction of timesteps once mixed with a stochastically braking HV population. We flag this because the same reduced-gap construction recurs across the mixed-autonomy CA literature and, to our knowledge, its collision safety in heterogeneous populations is not always explicitly checked.

The remainder of this paper is organized as follows: Section 2 describes the base model, a collision-safety pitfall in reduced-gap AV models, the four AV car-following rules and their shared safety architecture, the

simulation protocol, and the VDR slow-to-start rule together with two independent methods for detecting hysteresis. Section 3 presents the fundamental-diagram comparison across AV rules (Section 3.1) and the hysteresis-suppression results (Section 3.2). Section 4 independently validates the shared safety-floor architecture at $v_{\max} = 1$ against an exact disordered-exclusion-process solution. Finally, we summarize the paper in Section 5, where the implications and the limitations of the study are also discussed.

2. Model and methods

2.1. Base model

We simulate a single-lane, periodic (ring) NaSch lattice of L sites, each holding at most one vehicle. Vehicle i 's state is its position x_i and integer velocity $v_i \in \{0, \dots, v_{\max}\}$. Writing d_i for the center-to-center distance, in sites, from vehicle i to the vehicle immediately ahead, so that $d_i - 1$ is the number of empty sites between them (see Section 2.6), the standard update rule advances the system by

$$v_i \leftarrow \min(v_i + 1, v_{\max}) \quad (\text{acceleration}) \quad (1)$$

$$v_i \leftarrow \min(v_i, d_i - g_{\min}) \quad (\text{safety/gap constraint}) \quad (2)$$

$$v_i \leftarrow \max(v_i - 1, 0) \text{ w.p. } p \quad (\text{random slow-down}) \quad (3)$$

$$x_i \leftarrow x_i + v_i \quad (\text{movement}) \quad (4)$$

applied in parallel to every vehicle, where $g_{\min}$ is the minimum allowed gap. Standard NaSch (Nagel and Schreckenberg, 1992) is recovered with $g_{\min} = 1$ and a uniform p . A human-driven vehicle (HV) uses $v_{\max} = 5$ and $p_{\text{HV}} = 0.5$ in the fundamental-diagram experiments of Section 3.1 and in the theory comparison of Section 4; the slow-to-start experiments of Section 3.2 instead use the VDR randomization parameters given in Section 2.9. Every vehicle of every type uses $g_{\min} = 1$ in every production run reported in this paper; the only exception is the deliberately unsafe diagnostic of Section 2.3.

2.2. Baseline validation

Before extending this engine to mixed AV/HV traffic, we validate our own implementation against the classic homogeneous-HV result it is built on. Figure 1 shows the simulated fundamental diagram at $v_{\max} = 5$, $p = 0.5$, $L = 1000$, giving $J_{\max} = 0.328$ at $\rho_c = 0.075$, matching the qualitative shape, a

steep free-flow branch, a sharp kink, and a slowly decaying congested branch, of the classic triangular fundamental diagram of Nagel and Schreckenberg (1992).

Quantitative agreement was further verified directly against that paper for the identical parameters ($v_{\max} = 5$, $p = 0.5$): Sec. 5 of Ref. (Nagel and Schreckenberg, 1992) states explicitly that “our maximum of the flow is only 0.32 vehicles per timestep” ($J_{\max} = 0.32$ vs. our 0.328, a $\sim 2.5\%$ relative difference), and its Fig. 6, a zoomed, unambiguously labeled version of the same simulation data, shows the peak directly at $\rho_c \approx 0.08\text{--}0.09$, $J_{\max} \approx 0.32\text{--}0.33$. The small remaining difference in ρ_c is consistent with our smaller system size and shorter relaxation/averaging than that paper uses for its Figs. 4/6 ($L = 1000$, $T_{\text{relax}} = 1000$ here, versus $L = 10^4$, relaxation over $10 \times L$ steps, and averaging up to $10^5\text{--}10^6$ timesteps there): a bulk quantity such as $J_{\max}$ self-averages over many vehicles and is far less sensitive to this difference than the precise location of a sharp kink on a finite density grid. This agreement confirms the simulation engine itself, prior to the introduction of any AV-specific extension.

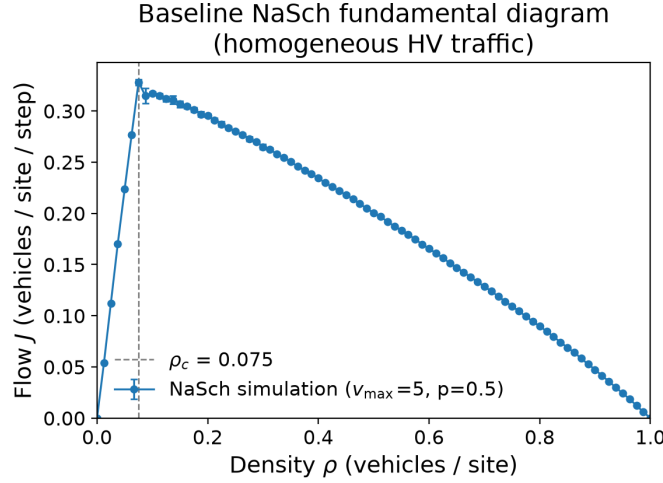


Figure 1: Baseline validation: simulated fundamental diagram for homogeneous HV traffic ($v_{\max} = 5$, $p = 0.5$), reproducing the classic NaSch triangular flow–density relation (Nagel and Schreckenberg, 1992; Chowdhury et al., 2000).

2.3. A collision-safety pitfall in reduced-gap AV models

A common way to represent an AV’s advantage over an HV is to let it tolerate a shorter headway, i.e., set $g_{\min,AV} < g_{\min,HV}$ in Eq. (2) (in the extreme, $g_{\min,AV} = 0$). Under the fully parallel update described above, this construction is collision-free only if the vehicle immediately ahead is guaranteed to move at least $g_{\min,HV} - g_{\min,AV}$ cells in any timestep in which it might be adjacent to a follower using the reduced margin, because Eq. (2) is evaluated against the leader’s *pre-update* position, not its realized post-update position. With $g_{\min,AV} = 0$, a follower may advance up to d_i cells, landing exactly on the leader’s pre-update site if the leader itself realizes $v = 0$ in that same timestep, for instance through an HV’s random slow-down or the leader’s own gap constraint.

We confirmed that this failure mode occurs in practice: instrumenting a homogeneous-AV population (all vehicles using $g_{\min} = 0$) shows zero collisions at any tested density, but a mixed population with as little as 20% AV, that is 80% HV, shows overlapping (duplicate-position) vehicle states in the large majority of measured timesteps at moderate density (e.g., $\rho = 0.1$, $r = 0.2$: 2515 of 3000 measured steps corrupted). The pitfall is thus invisible precisely in the homogeneous configuration in which it would most naturally be first tested, and manifests only once a heterogeneous population is introduced, exactly the regime this paper’s questions require. The mechanism, once stated, is elementary: under a fully synchronous update the follower computes its safe speed from the leader’s pre-update position, so a follower permitted to close to zero gap collides whenever the leader fails to move. We report it not as a subtle dynamical discovery but as a practical caution, since the reduced-gap construction recurs in the mixed-autonomy CA literature and the failure is silent: positions coincide without any error being raised. We therefore fix $g_{\min} = 1$ for *every* rule studied here, AV and HV alike; the AV/HV distinction is expressed entirely through the car-following dynamics described in Section 2.4, never through a shorter geometric margin. We have not found this specific mixed-population failure mode discussed in the mixed-autonomy cellular-automaton literature we surveyed, though we have not attempted an exhaustive search, and we report it here as a general caution for related model designs rather than as a claim of priority.

2.4. Four AV car-following rules

All four rules share the same architecture: each vehicle computes a *desired* velocity from its own car-following law, which is then subject to one shared

hard safety floor,

$$v_i \leftarrow \min(v_{i,\text{desired}}, d_i - 1), \quad (5)$$

identical in form to Eq. (2) with $g_{\min} = 1$, followed by random slow-down (Eq. 3, with $p = 0$ for every AV rule) and movement (Eq. 4). Only the first step, how $v_{i,\text{desired}}$ is computed, differs between rules.

2.4.1. Naive (baseline)

$v_{i,\text{desired}} = \min(v_i + 1, v_{\max})$, i.e. Eq. (1). This isolates the effect of removing random slow-down alone and serves as the point of comparison for the other three rules.

In the three continuous rules below, the spacing that enters the control law is the net bumper-to-bumper spacing

$$s_i \equiv d_i - 1, \quad (6)$$

not the center-to-center lattice distance d_i ; the reason, and the consequences of getting it wrong, are set out in Section 2.6.

2.4.2. IDM

Following Helbing (2001), Euler-discretized with $\Delta t = 1$ (one NaSch timestep):

$$v_{i,\text{desired}} = v_i + a \left[1 - \left(\frac{v_i}{v_0} \right)^\delta - \left(\frac{s^*(v_i, \Delta v_i)}{s_i} \right)^2 \right], \quad (7)$$

$$s^*(v_i, \Delta v_i) = s_0 + \max \left(0, T v_i + \frac{v_i \Delta v_i}{2\sqrt{ab}} \right), \quad (8)$$

where $\Delta v_i = v_i - v_{i+1}$ is the closing rate with the leader. We set $v_0 = v_{\max}$ (so all four rules share the same free-flow speed ceiling, isolating behavioral rather than desired-speed differences) and $s_0 = 0$ (the shared safety floor of Eq. 5 already reserves one cell, so an additional internal jam distance would double-count that margin). T , a , b are calibrated highway values verified directly from Treiber et al. (2000) ($a = 0.73 \text{ m/s}^2$, $b = 1.67 \text{ m/s}^2$, $T = 1.6 \text{ s}$, converted to cell/step units using the standard NaSch convention of one cell = 7.5 m and one timestep = 1 s), $\delta = 4$.

2.4.3. ACC and CACC

Following the experimentally calibrated control laws of Milanés and Shladover (2014) (verified directly against the published equations and coefficients), with gap error $e_i = s_i - hv_i$:

$$\text{ACC: } v_{i,\text{desired}} = v_i + k_1 e_i + k_2 (v_{i+1} - v_i), \quad k_1 = 0.23, \quad k_2 = 0.07, \quad h = 1.1 \text{ s}, \quad (9)$$

$$\text{CACC: } v_{i,\text{desired}} = v_i + k_p e_i + k_d (e_i - e_i^{\text{prev}}), \quad k_p = 0.45, \quad k_d = 0.25, \quad h = 0.6 \text{ s}. \quad (10)$$

The CACC control law requires the previous timestep’s gap error e_i^{prev} , carried as per-vehicle persistent state. Because k_1, k_2, k_p, k_d are purely inverse-time quantities and h is a time (and one NaSch timestep is by convention one second), these coefficients transfer to cell/step units unchanged. We note for completeness that our initial implementation assumed a single shared PD-control functional form for both ACC and CACC with a leader-acceleration feed-forward term for CACC; on verifying directly against Milanés and Shladover (2014), the two control laws are structurally different as given in Eqs. (9)–(10) (CACC is a velocity recursion without an explicit feed-forward term in the simplified traffic-flow model used here), while the numerical gain and time-headway values we had recalled were, once re-derived from the source, correct.

2.5. Velocity discretization and displacement rounding

Because IDM/ACC/CACC accelerations are frequently well under one cell/step² (IDM’s $a = 0.73/7.5 \approx 0.097$ cells/step²), rounding each vehicle’s velocity to an integer at every timestep silently discards this sub-unit progress every step; we found this produces permanent gridlock for a pure-IDM population under an integer-state implementation. We instead carry velocity as a persistent floating-point state across timesteps (exactly integer-valued for the naive/HV rule, genuinely fractional for IDM, ACC, and CACC) and round only the per-timestep *displacement* to an integer number of lattice cells; flux is measured from this integer displacement, not the continuous state.

Rounding the displacement to the nearest integer is not free of consequences of its own, and one of them is visible in our results. Any control law whose sustained equilibrium speed falls below half a cell per step is realized as no motion at all, and on a ring this is absorbing: nothing moves, the gaps stop changing, and the flow is reported as exactly zero. A homogeneous

IDM population reaches this condition on the congested branch, where the equilibrium speed v_{eq} , which is of order s_i/T , drops below 0.5 cells/step at $\rho \approx 0.65$ in these simulations; the IDM curve in Figure 2 is therefore identically zero for $\rho \gtrsim 0.7$, where the underlying continuous model would still be moving slowly. Readers should not interpret the high-density tail of the IDM curve as a property of IDM.

We have checked how far this reaches by rerunning the whole scan with an unbiased alternative, in which the fractional remainder of the displacement is carried forward to the next step instead of being discarded (the results are given immediately below and revisited in Finding 2). The gridlock disappears, as it must: the IDM flow at $\rho = 0.70$ becomes 0.074 rather than zero. The capacities move much less, because every capacity reported here is attained at $\rho \approx 0.07$ to 0.17, far from the affected region. The naive rule is unaffected to the last digit, since its velocity state is integer-valued throughout. IDM’s capacities become 0.334, 0.287, 0.301, 0.342, 0.412 across r , against 0.334, 0.276, 0.289, 0.339, 0.422 under rounding, so the non-monotonicity of Finding 1, its depth, and the recovery above the human-only baseline at $r = 1$ all survive. ACC changes by at most 0.006 up to $r = 0.75$ and by 0.018 at full penetration. The rule most affected is CACC, which moves by 0.015 at $r = 0.5$ and by 0.027 at $r = 0.75$, and falls from 0.830 to 0.763 at full penetration; we return to this in Finding 2.

2.6. Gap convention

One further convention must be stated explicitly. The quantity d_i computed on the lattice is the center-to-center distance from vehicle i to its leader, so it includes the cell the leader itself occupies; the net bumper-to-bumper spacing is $d_i - 1$. The shared safety floor, Eq. (5), is written in the net spacing. IDM, ACC, and CACC are likewise formulated in net spacing in their source papers, which is why Eqs. (7), (9) and (10) are written in s_i of Eq. (6) rather than in d_i . Supplying the gross distance instead hands those rules one extra cell, that is 7.5 m, of apparent room, which at the gaps typical of the densities near capacity is not a small perturbation. The first version of this preprint used the gross distance; Section 3.1 reports the capacity under both conventions so that the size of the difference is on the record.

2.7. Simulation protocol and AV/HV assignment

All simulations are on a periodic ring with parallel (fully synchronous) update and no passing, so the sequence of vehicle types around the ring is

quenched for the duration of a run. For a ring of L cells at density ρ we place $N = \text{round}(\rho L)$ vehicles; of these, $N_{\text{AV}} = \text{round}(rN)$ are designated AVs by drawing N_{AV} indices uniformly at random without replacement from the N vehicles, so that AVs are placed at uniformly random positions in the ring order. The draw is repeated independently for every seed, so that variation across seeds includes variation in the AV/HV sequence as well as in the stochastic dynamics. This matters because sequencing is known to affect hysteresis in mixed traffic (Jiang and Ahn, 2026); we do not control the sequence; we randomize over it, and the uncertainties reported below are therefore inclusive of it.

Table 1 gives the parameters of every experiment family reported in this paper. Random seeds are derived deterministically from a fixed base seed and the indices of the parameter point, so every figure and table can be regenerated exactly.

Table 1: Simulation protocol by experiment family, in the order the experiments appear in Sections 2.2, 3.1, 3.2 (VDR baseline, adiabatic ramp, penetration scan and standing composition) and 4. T_{relax} is discarded before measurement; T_{meas} is averaged over. HVs use $p_{\text{HV}} = 0.5$ in the baseline, AV-rule-comparison and theory families and the VDR parameters $p = 1/64$, $p_0 = 0.75$ in the other four; AVs are deterministic ($p = p_0 = 0$) in every family.

Experiment	L	density grid	r grid	T_{relax}	T_{meas}	seeds
Baseline FD	1000	81 pts in $[0, 1]$	n/a	1000	2000	8
AV rule comparison	1000	0.02–0.24 step 0.01, 0.25–0.90 step 0.05	$\{0, 0.25, 0.5, 0.75, 1\}$	1500	2000	16
VDR baseline	2000	22 pts in $[0.02, 0.45]$	0	3000	3000	3
Adiabatic ramp	1000	30 pts in $[0.02, 0.45]$	0	1500	1500	1
Penetration scan	1000, 2000, 4000, 8000	0.04–0.40 step 0.01	0–1 step 0.025	2000	4×1000	20
Standing composition	2000	0.15	0–1 step 0.1	2000	4000	10
Theory comparison	2000	25 pts	$\{0, 0.2, \dots, 1\}$	2000	2000	5

Flow is measured as the mean over the measurement window of the total displacement per timestep divided by L . For the rules whose velocity state is continuous (Section 2.4), the displacement used for this average is the integer number of cells actually advanced, not the continuous state velocity.

2.8. What the comparison does and does not isolate

The naive rule is native to the lattice, whereas IDM, ACC, and CACC are continuous-time control laws that have been discretized to run on it,

and each is then passed through the same hard safety floor. It is fair to ask whether the differences reported in Section 3.1 reflect the control laws themselves or the discretization and the shared constraint. They reflect the control laws *as embedded on this lattice under this safety architecture*, and we take that embedding to be the object of study rather than an incidental feature of it. We do not claim to measure a property of IDM or of CACC in the continuum, where those models are already well characterized; we claim to measure what happens when each is deployed in a mixed population under a common collision-safety constraint, which is the situation a CA study of mixed traffic actually models.

That the two can differ is not a hypothetical concern, and one of our results is precisely a case of it: Finding 2 in Section 3.1 reports that CACC’s distinguishing behavior is masked because its short desired headway makes its control law ask for more speed than the one-cell safety floor will grant, so that the floor rather than the control law sets its speed and on this lattice it becomes indistinguishable from the naive rule at high penetration. That is a statement about the embedding, not about CACC, and we label it as such. A reader interested in the continuum behavior of these controllers should not read Section 3.1 as evidence about it; a reader interested in what a lattice model of mixed traffic will report about them should.

2.9. *Slow-to-start rule and hysteresis detection*

To study metastability, we add the velocity-dependent-randomization (VDR) slow-to-start rule of Barlović et al. (1998): the random slow-down probability in Eq. (3) depends on whether a vehicle’s velocity at the start of the timestep was zero,

$$p(v) = \begin{cases} p_0 & v = 0 \\ p & v > 0 \end{cases}, \quad p_0 \geq p, \quad (11)$$

with HVs using Barlović et al.’s parameters ($p = 1/64$, $p_0 = 0.75$) and AVs deterministic regardless of velocity ($p = p_0 = 0$); both use the shared $g_{\min} = 1$ floor.

The choice $p = p_0 = 0$ is a strong idealization and deserves a word. It is the standard representation of an AV in the mixed-autonomy CA literature, and it is chosen here deliberately rather than for convenience: it isolates the removal of human stochasticity as the single difference between the two populations, so that any change in metastability can be attributed to that

removal and not to a competing change in desired headway or acceleration. It also makes the $r = 1$ limit exact rather than merely small, which provides the consistency check used in Section 3.2. A real automated vehicle is of course not noiseless, and an intermediate choice $0 < p_{AV} < p_{HV}$ would interpolate between the two populations studied here; since the suppression we report is continuous in r , we expect it to be continuous in the residual AV noise as well, but we have not tested this and it is left as future work.

We use two independent methods to detect hysteresis, both used by Barlović et al. (1998) to trace the same loop. *Method 1* (adiabatic ramp): starting from a low-density homogeneous configuration, density is increased by inserting vehicles into the continuing trajectory (no reset), tracing the upper (free-flow) branch; starting from a high-density jammed configuration, density is decreased by removing vehicles, tracing the lower (jammed) branch. *Method 2* (two initial conditions): flow is measured independently at each density from a fresh homogeneous initial condition (upper branch) and a fresh single-jam initial condition (lower branch). Method 1 is run only for the human-only ($r = 0$) baseline, since it is inherently sequential and expensive to repeat across many r ; Method 2, cheaper and embarrassingly parallel, is used for the full AV-fraction scan.

3. Simulation results

3.1. AV behavior rule comparison

Figure 2 shows the fundamental diagram at $r = 1$ (a homogeneous population under each rule); Figure 3 and Table 2 show maximum flow $J_{\max}(r)$ across AV fraction $r \in \{0, 0.25, 0.5, 0.75, 1.0\}$, mixed with a common HV background ($p_{HV} = 0.5$). Here and throughout, $J_{\max}(r)$ is the maximum, over a density grid refined to $\Delta\rho = 0.01$ near the low-density peak, of the flow averaged over 16 random initial conditions; the fine grid and seed averaging are needed because the IDM fundamental diagram peaks sharply at low density ($\rho \approx 0.07$ – 0.11 , depending on r), where a coarse grid or few seeds estimates the peak flow unreliably. As required by construction, the four rules agree at $r = 0$ (no AVs present, regardless of which AV rule is configured) to within about two combined standard errors, providing a built-in consistency check.

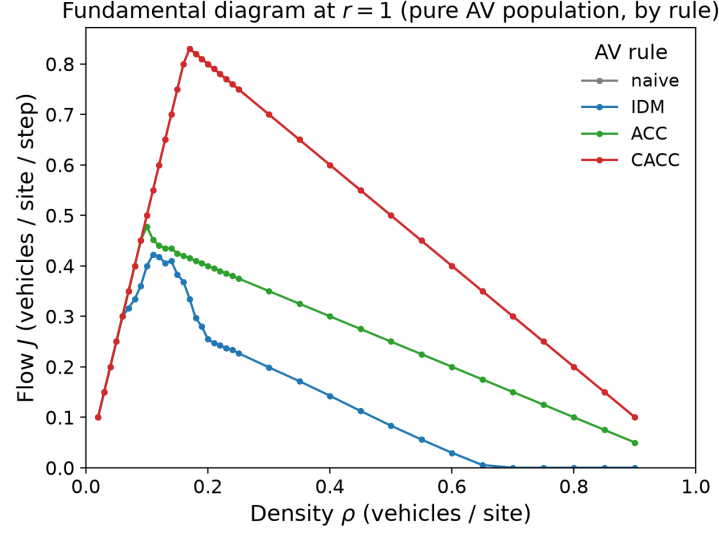


Figure 2: Fundamental diagram at $r = 1$ (homogeneous AV population) for each of the four car-following rules.

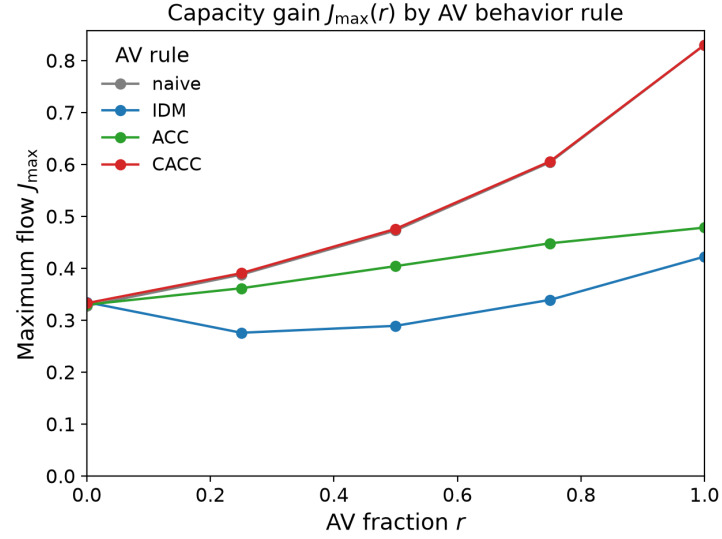


Figure 3: Maximum flow $J_{\max}$ as a function of AV fraction r , by AV car-following rule.

Table 2: Maximum flow $J_{\max}(r)$ by AV rule. Parenthesized figures are standard errors of the mean over the 16 seeds at the maximizing density. Each seed draws an independent AV/HV arrangement around the ring, so this scatter includes the effect of the quenched sequence. A zero entry means the standard error rounds to zero at the printed precision, which happens where the flow is set by the safety floor rather than by the stochastic dynamics.

r	naive	IDM	ACC	CACC
0.00	0.327 (0.002)	0.334 (0.003)	0.330 (0.002)	0.333 (0.002)
0.25	0.388 (0.002)	0.276 (0.004)	0.361 (0.003)	0.390 (0.003)
0.50	0.473 (0.002)	0.289 (0.004)	0.404 (0.003)	0.476 (0.002)
0.75	0.604 (0.003)	0.339 (0.006)	0.448 (0.000)	0.606 (0.002)
1.00	0.830 (0.000)	0.422 (0.000)	0.478 (0.006)	0.830 (0.000)

3.1.1. Sensitivity to the gap convention

Before turning to the findings, we record how much the gap convention of Section 2.6 matters, since the first version of this preprint used the gross center-to-center distance in the continuous rules. Table 3 repeats the scan under both conventions with identical seeds.

Table 3: Capacity $J_{\max}(r)$ under the gross (center-to-center) and net (bumper-to-bumper) gap conventions in the continuous rules. The naive rule does not use the gap in its own law and is unaffected by construction.

r	naive		IDM		ACC		CACC	
	gross	net	gross	net	gross	net	gross	net
0.00	0.327	0.327	0.334	0.334	0.330	0.330	0.333	0.333
0.25	0.388	0.388	0.268	0.276	0.379	0.361	0.390	0.390
0.50	0.473	0.473	0.265	0.289	0.439	0.404	0.476	0.476
0.75	0.604	0.604	0.269	0.339	0.519	0.448	0.609	0.606
1.00	0.830	0.830	0.275	0.422	0.611	0.478	0.830	0.830

The difference is not a rounding-level correction. Handing IDM one extra cell of apparent room lowers its capacity substantially, from 0.422 to 0.275 at $r = 1$, a reduction of 35%, which is the same comparison as the 53% gain obtained by correcting the convention; the mechanism is that the controller then persistently requests a speed that the safety floor must clip, and the resulting mismatch between the internal continuous state and the enforced

speed destabilizes the flow, collapsing it above $\rho \approx 0.07$. ACC moves the other way, its capacity falling from 0.611 to 0.478 at $r = 1$, by 22%, once it is given the correct, smaller spacing. CACC is essentially untouched, which is itself informative: as Finding 2 below explains, CACC’s requested headway is already short enough that the shared safety floor, not the control law, sets its speed, so a one-cell change in the spacing the control law sees does not reach the dynamics. All results reported from here on use the net convention. The net columns of Table 3 are the same runs as Table 2, sharing a seed stream, and agree with it exactly. Table 4 comes from a separate scan with an independent seed stream, so the same quantity can differ from the other two tables by of order the seed-to-seed scatter reported in Table 2; the IDM capacity at $r = 1$, for instance, appears as 0.422 there and 0.424 here.

Two findings depart from the simple, monotonic “more AVs, more capacity” picture.

3.1.2. Finding 1: an IDM-governed AV imposes a capacity penalty on mixed populations, and the penalty is a calibration effect

With the IDM parameters of Section 2.4 ($v_0 = v_{\max}$, $s_0 = 0$, $a = 0.73 \text{ m/s}^2$, $b = 1.67 \text{ m/s}^2$, $T = 1.6 \text{ s}$, values verified directly from Treiber et al. (2000), as is common practice when an IDM-based rule is assigned to an “AV” role), the capacity is not monotonic in r (Table 2). It begins at 0.334 for the all-HV population, falls to 0.276 at $r = 0.25$ and 0.289 at $r = 0.5$, then recovers to 0.339 at $r = 0.75$ and reaches 0.422 at full penetration. A homogeneous IDM population therefore carries more flow than a homogeneous human population, and a mixture of the two carries less than either: the deficit at $r = 0.25$ is about 17% relative to the human-only baseline. The four rules necessarily describe the same physical system at $r = 0$, and their tabulated values there (0.327, 0.334, 0.330, 0.333) differ only through the seed streams each consumes; we take their mean, 0.331, as the baseline for such comparisons.

This is worth stating carefully, because the first version of this preprint, which used the gross gap convention corrected in Section 3.1.1, reported instead that IDM lowers capacity at every penetration rate tested. That statement was an artifact of the convention. What survives the correction, and is arguably the more interesting statement, is that the penalty is specific to *mixed* populations rather than to automation as such.

The mechanism of the mixed-population penalty is that an IDM vehicle following a stochastically braking HV must accommodate closing-speed

fluctuations that its own modest actuation cannot absorb, so it falls back and leaves a gap that the following queue cannot use. Because $v_0 = v_{\max}$ and $s_0 = 0$, IDM’s own acceleration (Eq. 7) is never exactly zero except in the $d_i \rightarrow \infty$ limit: even a vehicle at $v_{\max}$ experiences residual deceleration from the interaction term at any finite gap, and this drag is amplified by the discrete lattice’s coarse velocity quantization acting on small closing-speed fluctuations, since a, b are numerically small in cell units.

To test whether this is a fixed property of IDM or an artifact of reusing a human-calibrated a, b for an AV role, we re-ran the same r -scan under three additional parameter sets: doubled acceleration/deceleration capability (“responsive”: $a = 1.46$, $b = 3.34 \text{ m/s}^2$), a shorter desired time headway alone (“short-headway”: $T = 0.8 \text{ s}$, a, b unchanged), and both combined (Table 4, Figure 4). Under the responsive calibration the mixed-population deficit disappears: at $r = 0.5$ the capacity is 0.344, above the human-only baseline of 0.328 measured in this same scan rather than below it, and at $r = 1$ it rises to 0.444. Shortening T alone, with a, b held at the weaker human-calibrated values, instead reduces capacity further, and does so most severely where the baseline calibration was strongest, at $r = 1$ (0.291 against the baseline calibration’s 0.424), because demanding a shorter gap without the actuation capability to maintain it safely is counterproductive; T and (a, b) interact rather than acting independently, as the combined set confirms: pairing the shorter headway with the stronger actuation recovers the gain and gives the highest capacity of the four calibrations at $r = 1$ (0.475). The correct statement of Finding 1 is therefore not that IDM reduces capacity as a fixed property, but that *reusing a human-driver-calibrated a, b for an AV role, without revisiting it for the AV’s presumably faster actuation, makes an IDM-governed AV depress the capacity of the mixed traffic it is deployed into, an effect that disappears once acceleration and deceleration capability are recalibrated, but only if the desired headway is not shortened faster than that capability can support.*

Table 4: IDM parameter sensitivity: $J_{\max}(r)$ under four calibrations.

Calibration	$a \text{ (m/s}^2\text{)}$	$b \text{ (m/s}^2\text{)}$	$T \text{ (s)}$	$J_{\max}(0)$	$J_{\max}(0.5)$	$J_{\max}(1)$
baseline (verified, Treiber et al., 2000)	0.73	1.67	1.6	0.328	0.288	0.424
responsive (2x a, b)	1.46	3.34	1.6	0.335	0.344	0.444
short-headway (lower T only)	0.73	1.67	0.8	0.329	0.278	0.291
combined (2x a, b + lower T)	1.46	3.34	0.8	0.332	0.346	0.475

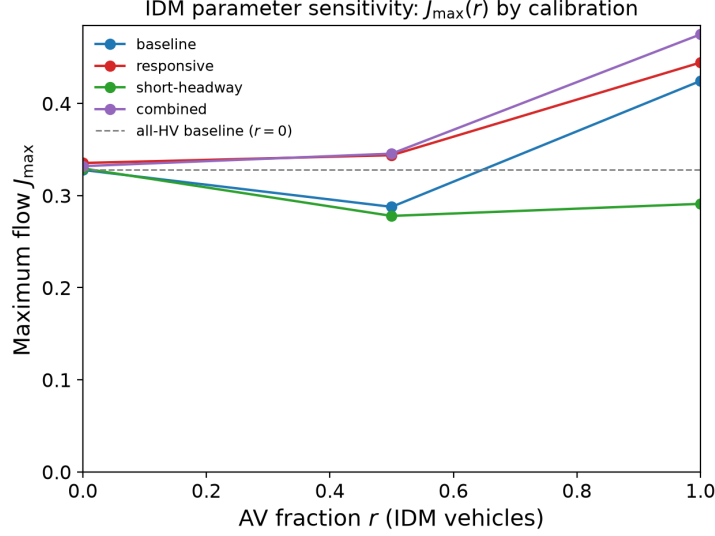


Figure 4: IDM parameter sensitivity: $J_{\max}(r)$ under four IDM acceleration/deceleration and time-headway calibrations.

3.1.3. Finding 2: CACC saturates against the shared safety floor

At $r = 1$, CACC's $J_{\max}$ (0.830) coincides with the naive rule's, while ACC's (0.478) is meaningfully lower. At CACC's own control-law fixed point ($e_i = e_i^{\text{prev}}$, so the derivative term vanishes), the implied equilibrium velocity is $v_{\text{eq}} = s_i/h$. With $h = 0.6$ s and the moderate gaps typical of the densities studied, v_{eq} substantially exceeds $v_{\max}$, so CACC's own control law is floor-bound (Eq. 5) at essentially every density tested; since the naive rule's acceleration ramp converges to the same $\min(v_{\max}, d_i - 1)$ fixed point given enough relaxation time, the two rules become numerically indistinguishable in steady state. The coincidence at $r = 1$ should not be read too literally, however. Under the unbiased displacement scheme discussed in Section 2.5, CACC at $r = 1$ gives 0.763 against the naive rule's unchanged 0.830, so part of the exact numerical agreement under rounding comes from fractional speeds being rounded up to the floor value rather than from saturation alone. What is robust is the ordering and its size: CACC remains close to the naive rule and far above ACC under both schemes. It is the ordering, not the coincidence, that supports the conclusion below.

ACC's smaller gains ($k_1 = 0.23$ vs. CACC's $k_p = 0.45$) and longer time headway ($h = 1.1$ s) leave it non-floor-bound more often within the mea-

surement window, producing a visibly different, lower curve. This indicates that control-law gains and time headways calibrated for continuous, meter-scale following distances in flowing real-world traffic can lose their distinguishing behavior once mapped onto a coarser cellular-automaton lattice at low-to-moderate density, a discretization-interaction effect, not a defect in the underlying verified control law.

3.2. VDR hysteresis suppression by AV penetration rate

Throughout this section the AVs use the naive deterministic rule of Section 2.4, with $p = p_0 = 0$. IDM, ACC, and CACC were not run under the VDR dynamics, so nothing below speaks to whether the suppression of hysteresis depends on the AV behavior rule; that combination is left as future work (Section 5). What this section establishes is the shape of the suppression for one AV rule.

Figure 5 validates the VDR implementation at $r = 0$ against Barlović et al.’s (1998) own result (two-initial-condition method): a clear hysteresis loop is reproduced, with the homogeneous-IC branch far above the jammed-IC branch in an intermediate density window, converging outside it. The supplementary adiabatic-ramp check (Figure 6) shows a weaker, noisier version of the same signal: inserting vehicles at random empty sites into an already-relaxed homogeneous trajectory introduces local perturbations that accumulate and nucleate jams earlier than a fresh restart would, so the ramp’s increasing branch does not sustain the same high-flow values as the two-IC method at comparable densities, though the decreasing branch does remain measurably above the increasing branch through most of the mid-density range. We rely on the two-IC method, Barlović et al.’s own primary approach, for the main AV-fraction result below.

Figure 7 illustrates the hysteresis loop at several AV fractions, using the coarse six-point scan of version 1 of this preprint. It is shown to convey the shape of the loop and its closure, not to support any quantitative claim; every loop width reported below comes from the fine scan described next. The loop width is defined as $\max_{\rho} (J_{\text{hom}}(\rho) - J_{\text{jam}}(\rho))$.

Resolving the shape of that width as a function of r requires more care than a coarse scan provides, and the first version of this preprint did not provide it. We therefore scan r in steps of 0.025 from 0 to 1, at densities spaced by 0.01 over $[0.04, 0.40]$, with twenty independent seeds per point and at four ring sizes $L \in \{1000, 2000, 4000, 8000\}$; the AV/HV placement around the ring is redrawn for each seed, so the reported scatter includes the

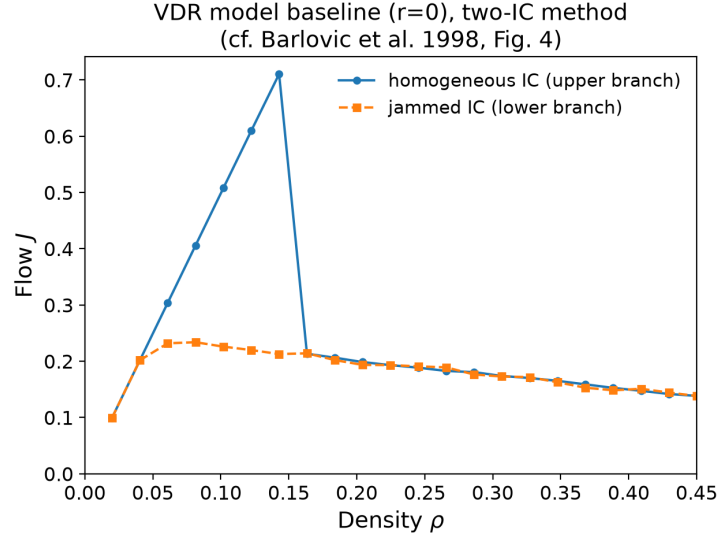


Figure 5: VDR model baseline ($r = 0$), two-initial-condition method, reproducing Barlović et al.'s (1998) Fig. 4.

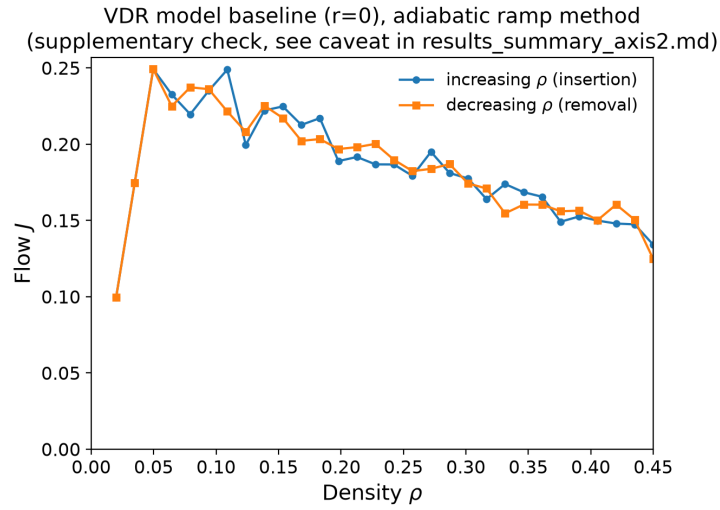


Figure 6: VDR model baseline ($r = 0$), adiabatic ramp method (supplementary check; see discussion of its weaker signal in the text).

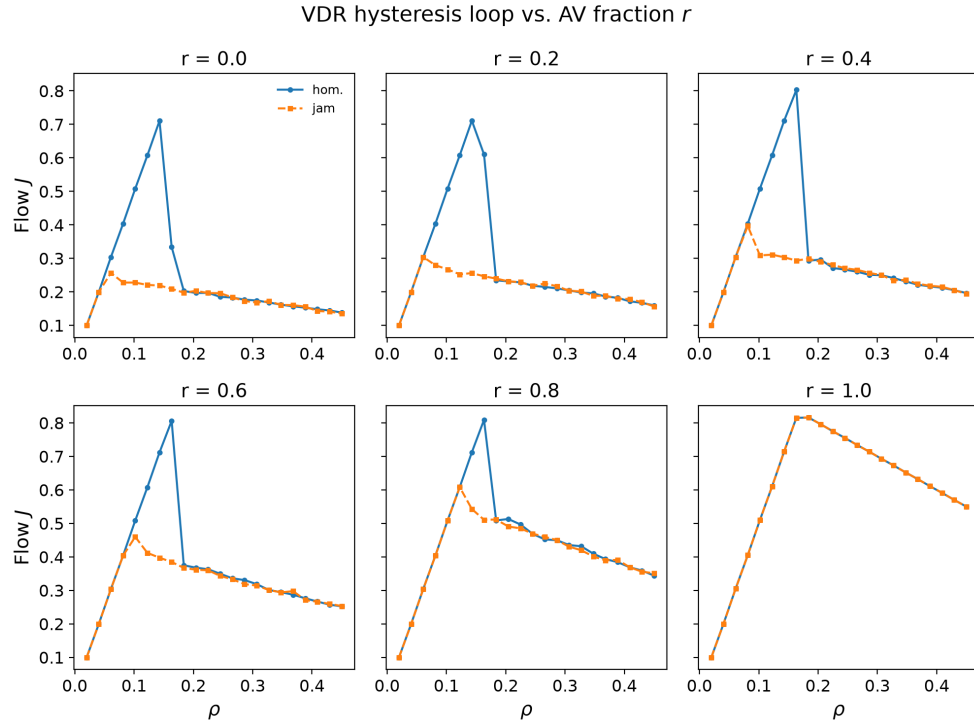


Figure 7: VDR hysteresis loop at several AV fractions r , from the coarse six-point scan of version 1 of this preprint ($L = 1000$, three seeds). Illustrative only; the quantitative loop widths of Table 5 come from the fine scan of Figure 8.

quenched AV/HV sequence, whose relevance in mixed traffic is emphasized by Jiang and Ahn (2026). Uncertainties are percentile intervals from two thousand bootstrap resamples over seeds. Flow is measured in four consecutive windows of 1000 steps after 2000 relaxation steps, so that any decay of the metastable branch during measurement would be visible. Loop widths from the first and the last window agree to within 0.056 in absolute terms at every r and L . There is a slight systematic narrowing between the two, a mean decrease of 0.007 with 108 of the 164 measured points decreasing, which is consistent with a slow decay of the metastable branch. It is an order of magnitude smaller than the variation in W across r that Finding 3 concerns, and, being present at all penetration rates rather than concentrated near any particular one, it does not affect the shape of the curve.

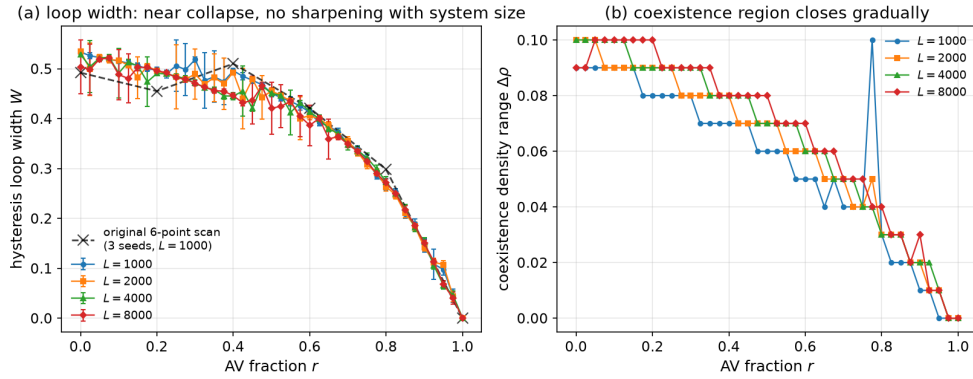


Figure 8: Hysteresis loop width (a) and two-branch coexistence density range (b) against AV fraction, at four ring sizes. Error bars in (a) are 95% bootstrap intervals over seeds. The dashed curve in (a) is the coarse six-point scan of version 1 of this preprint, whose apparent plateau and rise near $r = 0.4$ lie within the scatter of a three-seed measurement. The single outlying point near $r = 0.78$ in (b) is an artifact of applying a hard cutoff to a noisy branch separation.

3.2.1. Finding 3: hysteresis suppression is a continuous crossover, not a threshold

The loop width falls monotonically, within uncertainty, from $W \approx 0.53$ at $r = 0$ to zero at $r = 1$, and the fall accelerates at high penetration. It is tempting, and it was our earlier reading, to call this threshold-like. That reading is not supported, for three reasons, of which the first two concern finite-size behavior.

Table 5: Hysteresis loop width by AV fraction at four ring sizes, and the density range $\Delta\rho$ over which the two branches remain separated by more than 0.02 (at $L = 8000$).

r	loop width W				$\Delta\rho$
	$L = 1000$	$L = 2000$	$L = 4000$	$L = 8000$	
0.00	0.533	0.535	0.528	0.504	0.09
0.20	0.497	0.495	0.490	0.496	0.10
0.40	0.496	0.493	0.445	0.446	0.08
0.60	0.413	0.406	0.414	0.387	0.07
0.80	0.268	0.264	0.273	0.271	0.04
0.90	0.146	0.142	0.153	0.150	0.03
0.95	0.098	0.107	0.066	0.068	0.01
1.00	0.000	0.000	0.000	0.000	0.00

First, the curves for the four ring sizes very nearly collapse. Over a factor of eight in L the spread across ring sizes never exceeds 0.058 in loop width, against a loop width of 0.53 at $r = 0$ (Table 5). The collapse is not exact, and we state the residual in full because the shape of this curve is the result at issue: the spread exceeds five percent of the local width at 22 of the 40 penetrations at which the loop is open, with the largest relative spreads occurring at $r \geq 0.95$, where the loop is nearly closed and the width itself is small. There is also a weak but systematic tendency for the larger rings to give slightly smaller widths: the r -averaged width falls from 0.388 at $L = 1000$ to 0.375 at $L = 8000$, $L = 8000$ lies below $L = 1000$ at 30 of the 41 penetrations tested, and the 95% intervals of those two sizes separate at five of them. This residual finite-size effect is small, and, more to the point, it has the wrong sign to indicate a threshold: a sharpening transition would make the loop width more extreme as L grows in the transition region, not uniformly smaller. The loop width is therefore a well-defined bulk quantity here up to a mild finite-size correction, and there is no size-dependent sharpening.

Second, and following from this, the descent does not steepen as the system grows. The steepest slope $|dW/dr|$ anywhere in the scan is 1.95, 2.15, 1.86 and 1.64 for $L = 1000, 2000, 4000$ and 8000 respectively, attained in every case at $r \geq 0.925$: flat, if anything mildly decreasing. A genuine transition would steepen without bound as L grows, approaching a discontinuity in the limit. Fitting competing functional forms bears this out. A continu-

ous piecewise-linear form with a free breakpoint does win on AIC over linear and power-law forms at every ring size, confirming that the decline is genuinely nonlinear, but the fitted breakpoint drifts upward with system size ($r_c = 0.672, 0.673, 0.697, 0.716$ for the four sizes) instead of converging on a fixed value, which is the signature of a crossover rather than of a critical point.

Third, the two-branch coexistence region closes gradually rather than abruptly (Figure 8b). The density range over which the homogeneous-start and jammed-start branches remain separated is 0.09 at $r = 0$, rises by one grid step to 0.10 by $r = 0.05$ and holds there to $r = 0.2$, and thereafter falls to 0.04 by $r = 0.8$ and to zero at $r = 1$, with one further one-step rise at $r = 0.9$. Both rises are single steps of the 0.01 density grid and should not be read as structure. This observable has the advantage of not depending on where in the density range the maximum separation happens to fall, which makes it a useful cross-check on the loop width, but its resolution is coarse and we do not rest the argument on it alone.

We therefore conclude that AV penetration closes the slow-to-start coexistence continuously, and that no critical penetration rate exists in this model. The vanishing at $r = 1$ remains exact rather than approximate: an all-AV population has $p = p_0 = 0$ uniformly, so the VDR rule reduces identically to the deterministic NaSch model, which cannot produce metastability by construction. That limit is a built-in consistency check on the scan rather than a separate finding, and, as it follows by construction, it cannot by itself establish the shape of the approach to it, which is why the fine scan was needed.

3.2.2. The AV share of standing vehicles tracks the penetration rate

Version 1 of this preprint explained the shape of the suppression curve by asserting that AVs must come to dominate the population of standing vehicles before the slow-to-start asymmetry is diluted. That assertion is directly measurable, and the measurement does not support it. We measure it at $\rho = 0.15$, which lies inside the coexistence window at every penetration rate measured in this experiment for which a coexistence window exists, starting from the jammed initial condition and averaging over ten seeds and four thousand measured steps. The AV share of vehicles with $v = 0$ equals r to within 5×10^{-4} at every penetration rate tested, and the mean number of consecutive steps a vehicle spends standing is the same for the two types: 329.0 steps for AVs against 329.1 for HVs at $r = 0.5$, and 45.7 against 45.0

at $r = 0.9$. Standing vehicles are neither enriched nor depleted in AVs.

The reason is that inside a compact jam a vehicle is held stationary by the gap constraint, which is shared by both types, and not by its own randomization: the slow-to-start rule can only act on a vehicle that already has room to move, that is, at the downstream jam front. What does change with r is the amount of standing itself. The standing fraction on the jammed branch falls smoothly from 0.715 at $r = 0$ through 0.539 at $r = 0.5$ to 0.088 at $r = 0.9$, and the mean standing time, pooled over both types, falls from 733 steps to 329 and then to 46 steps over the same range. The suppression of hysteresis follows this smooth decline, which is consistent with the absence of a threshold reported above and offers no mechanism by which one would arise.

4. Analytical validation at $v_{\max} = 1$

The simulation results above (Section 3) all use $v_{\max} = 5$, for which no exact analytical solution of the mixed AV/HV model is available. As an independent check on the modeling framework itself, specifically the naive rule and shared safety-floor architecture of Section 2.4, we additionally verify it against an exact result at $v_{\max} = 1$, where the model reduces to a special case of a well-studied class of exactly solvable exclusion processes.

4.1. Connection to disordered exclusion processes

At $v_{\max} = 1$, our mixed model (vehicle i moves one site forward if the site ahead is empty, with probability $q_\tau = 1 - p_\tau$ depending on type $\tau \in \{\text{AV}, \text{HV}\}$, $p_{\text{AV}} = 0$, $p_{\text{HV}} > 0$) is an instance of the totally asymmetric exclusion process (TASEP) on a ring with *particle-wise disordered hopping probabilities* under fully parallel (synchronous) updating. This disordered-TASEP class was solved exactly by Evans (1997) for an arbitrary distribution of per-particle hopping probabilities q_μ (in Evans’s notation, particle μ hops with probability p_μ ; our q_τ plays the same role). Our two-type case is thus the special case of Evans’s (1997) result restricted to a two-point distribution ($q = 1$ for AV, $q = q_{\text{HV}} < 1$ for HV, in fractions r and $1 - r$). The exact stationary gap-distribution weights (Evans, 1997, Eq. 4), the current/velocity self-consistency equation (Evans, 1997, Eq. 17), and the fact underlying both, that no two adjacent particles on a no-passing ring can sustain different long-run average velocities, since a persistent gap between differing average speeds would grow or shrink without bound on a bounded

ring, are all already established there for the fully general disordered case. We restate the result below in the car-oriented mean-field theory (COMF) language used in the traffic-CA literature (Schadschneider and Schreckenberg, 1997) and specialize it to the two-type AV-fraction parametrization relevant to mixed autonomous/human traffic, for the reader's convenience and as a self-contained basis for the validation that follows, and we validate the resulting flux-density relation against our own simulation. We make no claim of priority for either step.

4.2. Two-type closed form and density constraint

Since $q_{AV} = 1$, adjacent vehicles' shared move probability (J/ρ , the mean velocity, equivalently Evans's velocity v , here written g^* to match COMF notation) lets the AV gap distribution be read off directly from Evans's Eq. 4: $P_0^{AV} = 1 - g^*$, $P_1^{AV} = g^*$, $P_n^{AV} = 0$ for $n \geq 2$ (a deterministic follower's gap is supported only on $\{0, 1\}$), giving mean occupied width

$$\mathbb{E}[n + 1 \mid AV] = 1 + g^*. \quad (12)$$

For HV, the same equation gives the geometric tail $P_n^{HV} = (P_0^{HV}/p_{HV}) z_{HV}^n$ ($n \geq 1$), with $z_{HV} = p_{HV}g^*/(q_{HV}(1 - g^*))$ and $P_0^{HV} = (1 - g^*) - (p_{HV}/q_{HV})g^*$, where z_{HV} is the ratio governing the geometric tail and is not a density, giving

$$\mathbb{E}[n + 1 \mid HV] = P_0^{HV} \left[1 + \frac{1}{p_{HV}} \left(\frac{1}{(1 - z_{HV})^2} - 1 \right) \right]. \quad (13)$$

Since every vehicle occupies $n + 1$ sites, the density ρ satisfies

$$\frac{1}{\rho} = r \mathbb{E}[n + 1 \mid AV](g^*) + (1 - r) \mathbb{E}[n + 1 \mid HV](g^*), \quad (14)$$

a single scalar equation solved numerically for g^* (bracketed root-finding restricted to $g^* \in (0, q_{HV})$ whenever $r < 1$, since an HV's own move probability cannot exceed q_{HV}). At $r = 0$ this reduces exactly to the known single-species $v_{\max} = 1$ solution (Schadschneider and Schreckenberg, 1997); at $r = 1$ it reduces exactly to the standard deterministic triangular relation $J = \min(\rho, 1 - \rho)$.

4.3. Validation against simulation

Figure 9 and Figure 10 compare Eq. (14) (solved for $J = \rho g^*$) against direct simulation at $v_{\max} = 1$, $p_{HV} = 0.5$, across $r \in \{0, 0.2, 0.4, 0.6, 0.8, 1.0\}$.

Maximum relative error is below 0.53% at every r tested (mean below 0.2%), consistent with finite-size/finite-sampling noise rather than a systematic deviation. This confirms that the shared safety-floor architecture used throughout this paper (Section 2.4) behaves exactly as the underlying exclusion-process theory predicts in the one regime where an exact comparison is possible, supporting the correctness of the simulation framework used for the $v_{\max} = 5$ results in Section 3.

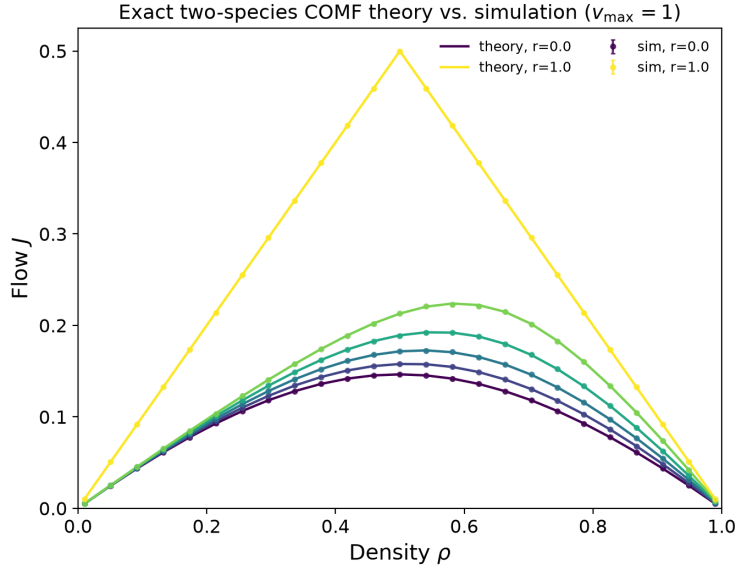


Figure 9: Exact two-type theory (solid curves) vs. simulation (points with error bars) at $v_{\max} = 1$, across AV fraction r .

5. Discussion and conclusions

Taken together, these results argue that “AV penetration improves traffic flow” is not a single, model-independent statement. Both the magnitude (Section 3.1, Table 2) and even the direction of the capacity effect (the IDM finding) depend on which AV car-following model is assumed and on how it is calibrated, and the hysteresis-suppression benefit (Section 3.2), measured here for the naive deterministic AV rule only, accrues continuously with penetration rather than at a threshold, tracking the smooth decline in how much standing the traffic sustains.

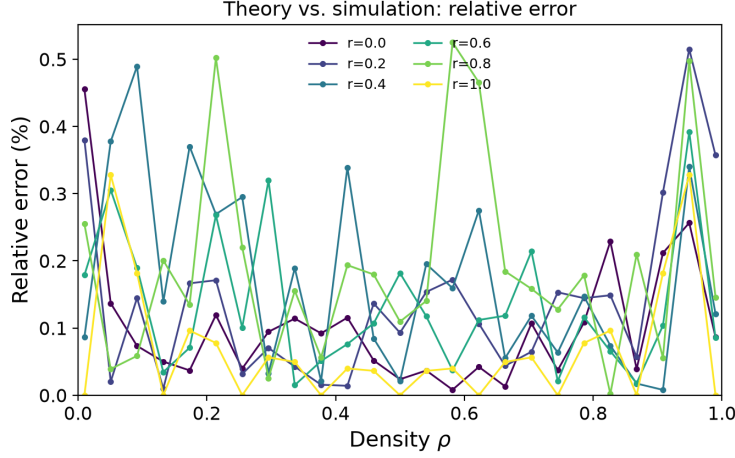


Figure 10: Relative error between theory and simulation, by density and AV fraction r .

Note that the exact vanishing of hysteresis at $r = 1$ and the no-passing assumption underlying Section 4 share a common setting: on a single-lane ring with no passing, no two adjacent vehicles can sustain different long-run average velocities, since a persistent gap between differing average speeds would grow or shrink without bound on a bounded ring. The CACC saturation of Finding 2 is a separate effect, arising from the one-cell resolution of the shared safety floor rather than from this argument.

Limitations.. All results are for a single lane with no lane changing, a simplification we adopt for three reasons rather than for convenience alone. First, it preserves direct comparability with the mixed-autonomy NaSch literature cited in Section 1, which is itself single-lane in almost every case. Second, and more fundamentally, the exact result of Section 4 depends on the no-passing assumption in an essential way: the shared-velocity argument underlying both Evans’s (1997) disordered-exclusion-process solution and our hysteresis-suppression result (Finding 3, Section 3.2) requires that a vehicle’s identity, and hence its long-run average velocity, never changes; once passing is permitted this no longer holds, and no analogous exact solution is available. Third, no lane changing is the more defensible approximation specifically in the congested, jam-forming regime that both the hysteresis (Section 3.2) and analytical-validation (Section 4) analyses target, since gaps wide enough to support a lane change become increasingly rare as density approaches ρ_c . Note that the same assumption is correspondingly harder to defend in the

free-flow regime, where a real vehicle could overtake a slower leader rather than being held behind it indefinitely; the model may therefore underestimate free-flow capacity, particularly for rule combinations in which a single slow leader suppresses an entire following queue (e.g., IDM at low r , Finding 1) that a multi-lane road would instead allow to pass. It has been put to us that Finding 1 is, for this reason, largely an artifact of the one-dimensional geometry: the capacity shortfall under human-calibrated IDM is the familiar slow-vehicle bottleneck, and a multi-lane road would let the queue behind a slow AV disperse. We accept the force of this objection and do not claim that a real, multi-lane mixed-traffic stream would lose capacity in the way reported here. What the single-lane result does establish is narrower and still worth stating: within a fixed geometry, holding the safety architecture constant, the sign of the capacity effect is set by the calibration of the following rule rather than by the presence of automation as such. We further speculate that the accelerating part of the hysteresis-suppression curve of Finding 3 would move toward lower AV fractions once lane changing lets AVs dissolve HV-induced jams more efficiently, but this remains untested. Extending the present framework to multiple lanes, and re-examining whether the rule-dependence and hysteresis-suppression findings above survive once passing is allowed, remains an important direction for future work.

The ACC/CACC parameters are one literature-consistent choice (Section 2.4); we did not re-run a corresponding parameter-sensitivity check for them, although the gap-convention comparison of Section 3.1.1 is informative in the same direction, showing ACC to be strongly sensitive to the spacing it is given and CACC to be almost entirely insensitive to it; the CACC/naive equivalence at $r = 1$ (Finding 2) is therefore not a demonstration that CACC’s control law is inherently insensitive to calibration, only that the one calibration tested here is floor-bound, and it remains open whether Finding 2 is similarly calibration-dependent or a more robust consequence of the shared safety floor; this is a natural next step. For IDM, we already confirmed (Section 3.1, Table 4) that the qualitative sign of the capacity effect is calibration-dependent, not fixed, so Finding 1 should be read as a statement about the sensitivity of IDM-based AV rules to their acceleration/deceleration calibration rather than a fixed property of IDM. The slow-to-start (Section 3.2) and car-following-rule (Section 3.1) mechanisms were also not yet combined: we did not test whether IDM, ACC, or CACC’s smoother, non-random deceleration dynamics interact differently with VDR-induced hysteresis than the naive deterministic rule tested in Section 3.2.

Implications for AV deployment.. We now consider the practical implications of these results for near-term AV deployment. Because the suppression of slow-to-start hysteresis is continuous rather than threshold-like (Finding 3, Section 3.2), there is no penetration rate that must be reached before any benefit appears; equally, there is no point at which a disproportionate benefit is unlocked. At the low penetration rates of current and near-term deployment the hysteresis benefit is correspondingly small, so the dominant lever available to planners remains the capacity effect of Section 3.1. We note that our result is specific to slow-to-start metastability in this model rather than to traffic instability in general, and that it does not speak to whether threshold behavior arises in other measures of automation benefit. Our results indicate that this capacity effect is not guaranteed to be positive: a fleet of AVs governed by an IDM-type rule calibrated on human driving data, without recalibrating acceleration and deceleration capability for the AV’s own actuation, may reduce rather than improve flow relative to human-only traffic during the mixed phase of deployment, even though the same fleet would outperform human traffic once it displaced it entirely (Finding 1). The deficit is therefore worst precisely in the transitional regime that any deployment must pass through. The reduced-gap collision-safety pitfall identified in Section 2.3 is a similarly concrete caution: a control law verified safe in homogeneous AV testing is not thereby verified safe in the mixed populations that any real deployment will pass through en route to full penetration. We expect these calibration- and safety-related cautions to be at least as relevant at higher, future AV penetration rates as at present.

Summary.. We have compared four AV car-following rules, a simple deterministic rule, IDM, ACC, and CACC, on the same NaSch lattice under a shared, collision-safe architecture, and have shown that AV penetration’s effect on capacity is strongly rule-dependent. This includes a case (IDM, under a human-driver-calibrated acceleration/deceleration setting) in which capacity is non-monotonic in penetration, so that mixed traffic carries less flow than either a purely human or a purely automated population, an effect that disappears under a more AV-appropriate recalibration and is therefore a calibration sensitivity rather than a fixed property of IDM, and a separate case (CACC) in which a literature-verified control law’s distinguishing behavior is masked by discretization-scale saturation against the shared safety floor. We have also found, independently, that AV penetration suppresses slow-to-start-induced hysteresis continuously rather than at a threshold: on

a fine penetration grid, at four ring sizes spanning a factor of eight, the loop width decreases monotonically within uncertainty, the decline accelerates at high penetration without the descent sharpening as the system grows, and the loop closes exactly at full penetration. This was tested for the naive deterministic AV rule alone. We have verified by direct measurement that the AV share of standing vehicles equals the penetration rate and that standing times do not differ by vehicle type, which removes the mechanism we had previously proposed for this curve and leaves a smooth decline in the amount of standing as the operative quantity. This corrects the corresponding claim in the first version of this preprint. Along the way, we have identified and corrected a collision-safety pitfall in reduced-gap AV models that is invisible in homogeneous populations and manifests only once mixed with stochastically braking vehicles; we believe this methodological point is relevant beyond the present study. Finally, we have validated the shared safety-floor architecture underlying all of the above against an exact result at $v_{\max} = 1$: our model reduces there to a special case of an already-known, exactly solvable disordered exclusion process (Evans, 1997), expressed here in car-oriented mean-field language (Schadschneider and Schreckenberg, 1997) and specialized to the AV-fraction parametrization, with simulation matching this exact prediction to within 0.53% across all penetration rates tested.

As discussed above, the rule-dependence of the capacity effect can be traced to how each behavior model interacts with the shared, collision-safe safety floor. The hysteresis result is of a different kind and should not be read as a rule-dependence claim: it was obtained for the naive deterministic AV rule alone, and it follows the smooth decline in the amount of standing that the traffic sustains as r grows. A single AV penetration rate cannot therefore be assigned a single capacity benefit independent of the behavior model in question; whether the same is true of the stability benefit is a question this paper leaves open.

CRediT authorship contribution statement

Mitsuyoshi Yagyū: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The simulation code, analysis scripts, and result data supporting the findings of this study are available from the corresponding author upon reasonable request.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- van Arem, B., van Driel, C.J.G., Visser, R., 2006. The impact of cooperative adaptive cruise control on traffic-flow characteristics. *IEEE Transactions on Intelligent Transportation Systems* 7, 429–436. doi:10.1109/TITS.2006.884615.
- Barlović, R., Santen, L., Schadschneider, A., Schreckenberg, M., 1998. Metastable states in cellular automata for traffic flow. *European Physical Journal B* 5, 793–800. doi:10.1007/s100510050504. arXiv:cond-mat/9804170.
- Chowdhury, D., Santen, L., Schadschneider, A., 2000. Statistical physics of vehicular traffic and some related systems. *Physics Reports* 329, 199–329. doi:10.1016/S0370-1573(99)00117-9.
- Evans, M.R., 1997. Exact steady states of disordered hopping particle models with parallel and ordered sequential dynamics. *Journal of Physics A: Mathematical and General* 30, 5669–5685. doi:10.1088/0305-4470/30/16/011. arXiv:cond-mat/9705006.

- Guo, X.Y., Zhang, G., Jia, A.F., 2023. Study on mixed traffic of autonomous vehicles and human-driven vehicles with different cyber interaction approaches. *Vehicular Communications* 39, 100550. doi:10.1016/j.vehcom.2022.100550.
- Helbing, D., 2001. Traffic and related self-driven many-particle systems. *Reviews of Modern Physics* 73, 1067–1141. doi:10.1103/RevModPhys.73.1067.
- Jiang, J., Ahn, S., 2026. Stochastic and dynamic fundamental diagram for mixed traffic. *arXiv preprint ArXiv:2604.24771*.
- Jiang, Y., Wang, S., Yao, Z., Zhao, B., Wang, Y., 2021. A cellular automata model for mixed traffic flow considering the driving behavior of connected automated vehicle platoons. *Physica A: Statistical Mechanics and its Applications* 582, 126262. doi:10.1016/j.physa.2021.126262.
- Kinjo, K., Tomoeda, A., 2025. Comparison of gap-based and flow-based control strategies using a new controlled stochastic cellular automaton model for traffic flow. *Transportation Letters* 17, 1171–1181. doi:10.1080/19427867.2024.2417150.
- Kumar, P., Chakraborty, P., Peruru, S.S., 2026. Macroscopic characteristics of mixed traffic flow with deep reinforcement learning based automated and human-driven vehicles. *arXiv preprint ArXiv:2603.25328*.
- Ma, L., Qu, S., Ren, J., Zhang, X., 2023. Mixed traffic flow of human-driven vehicles and connected autonomous vehicles: string stability and fundamental diagram. *Mathematical Biosciences and Engineering* 20, 2280–2295. doi:10.3934/mbe.2023107.
- Maerivoet, S., De Moor, B., 2004. Non-concave fundamental diagrams and phase transitions in a stochastic traffic cellular automaton. *European Physical Journal B* 42, 131–140. doi:10.1140/epjb/e2004-00365-8.
- Milanés, V., Shladover, S.E., 2014. Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. *Transportation Research Part C: Emerging Technologies* 48, 285–300. doi:10.1016/j.trc.2014.09.001.

- Nagel, K., Schreckenberg, M., 1992. A cellular automaton model for freeway traffic. *Journal de Physique I* 2, 2221–2229. doi:10.1051/jp1:1992277.
- Poudel, B., Li, W., Heaslip, K., 2024. EnduRL: enhancing safety, stability, and efficiency of mixed traffic under real-world perturbations via reinforcement learning, in: 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 13671–13678. doi:10.1109/IROS58592.2024.10802689. arXiv:2311.12261.
- Schadschneider, A., Schreckenberg, M., 1997. Car-oriented mean-field theory for traffic flow models. *Journal of Physics A: Mathematical and General* 30, L69–L75. doi:10.1088/0305-4470/30/4/005.
- Treiber, M., Hennecke, A., Helbing, D., 2000. Congested traffic states in empirical observations and microscopic simulations. *Physical Review E* 62, 1805–1824. doi:10.1103/PhysRevE.62.1805. arXiv:cond-mat/0002177.