From Microscopic Interactions to Network Congestion: A Multi-Scale Study through Simulation, Reinforcement Learning, Game-Theoretic Routing, and Optimal Transport

Adam Gassem[1], Mayeul Lambert[1], Amine Maazizi[1], Sammy Oukaci[1], and Ewerthon Melzani[1]

[1]MVA Master, ENS Paris-Saclay
{adam.gassem, mayeul.lambert, amine.maazizi, sammy.oukaci, ewerthon.melzani}@ensta.fr

August 18, 2026

Abstract

Traffic congestion emerges from local driver interactions: individual decisions about speed, headway, and route choice collectively produce free flow, stop-and-go waves, phantom jams, and gridlock that no single driver intends. This paper studies these phenomena across distinct yet interconnected modeling scales.

We first build a unified microscopic benchmark on a periodic ring road, comparing the Nagel–Schreckenberg cellular automaton [15], the Krauss stochastic safe-speed model [10], and the Intelligent Driver Model [20] under identical conditions. Our shared observer suite covers fundamental diagrams, jam morphology, spatiotemporal slow-car maps, velocity variance, and safety indicators including time-to-collision and gap-zero events.

We then evaluate local traffic control through reinforcement learning in a mixed-autonomy setting. A single policy is trained on one microscopic model and tested across all three, introducing controlled vehicles at increasing penetration rates from a single agent up to a fraction of the total fleet. We study how the number of controlled vehicles and their observation radius jointly determine the policy's ability to dissipate stop-and-go waves and recover aggregate flow, and whether a policy trained on one model transfers to the others without retraining.

To connect these microscopic instabilities to strategic network effects, we formalize road traffic as a congestion game and study Nash and Wardrop equilibria [17, 23], the Beckmann potential [2], and the Price of Anarchy [18]. We simulate four routing strategies on a grid network, from shortest-path to random-walk agents, identifying a demand threshold beyond which Braess's paradox [3] emerges. This shows that congestion is not only a dynamical instability but also a coordination failure sustained by selfish routing.

Finally, to unify these discrete agent behaviors within a rigorous continuum, we introduce a macroscopic graph-density optimal-transport model. Grounded in Mean Field Game theory [12] and Wasserstein gradient flows [1, 8], this multi-commodity framework models heterogeneous vehicle groups through a thermodynamic free energy functional. By balancing local entropy, destination potentials, and nonlocal congestion interactions, the continuum model organically reproduces the concave envelope of microscopic fundamental diagrams. Furthermore, it accurately predicts emergent network-scale phenomena, including capacity hysteresis and bidirectional phase-locking, while inherently resisting the systemic over-concentration that drives macroscopic coordination failures.

Code for the simulations and experiments is available at https://github.com/Mayeul84/interactions_road_traffic.

1. Introduction

Traffic congestion is one of the clearest examples of a complex system. Drivers react only to what is immediately around them, yet their collective behavior produces organized macroscopic states that nobody individually coordinates. Free flow, synchronized traffic, stop and go waves, and gridlock all arise from the same simple local rules repeated across hundreds of vehicles. Ring-road experiments [19] make this particularly transparent: congestion can appear with no bottleneck, no accident, and no external disturbance. It is a purely endogenous consequence of how vehicles interact.

Microscopic traffic models are the standard tool for studying these mechanisms in a controlled way. The Nagel-Schreckenberg (NaSch) model uses discrete space and time with stochastic braking to capture the exclusion and randomness present in real traffic [15]. The Krauss model works in continuous velocity space while enforcing an explicit safety constraint at each step [10]. The Intelligent Driver Model (IDM) encodes a smooth nonlinear response to desired gaps, time headways, and approach rates [20]. These three models represent genuinely different modeling philosophies but are usually studied separately. One of the goals of this paper is to compare them directly under a shared experimental protocol to identify which congestion phenomena are robust across models and which are specific to a given formulation.

Beyond model comparison, we ask whether congestion can be mitigated through local interventions. Motivated by the mixed-autonomy traffic literature [11, 22, 24], we train a reinforcement learning agent that controls a small number of vehicles on the ring road and measure the effect on global traffic. We introduce controlled vehicles progressively, from a single agent up to a significant fraction of the fleet, to understand how penetration rate shapes the outcome. A key question here is also informational. We investigate how much the agent needs to observe to act effectively. We test policies trained with access to global traffic state as well as policies limited to local sensing, to understand the gap between what is theoretically achievable and what a real deployed vehicle could do. We also test whether a policy trained on one microscopic model transfers to the others without retraining, which speaks to the practical robustness of learned controllers.

A separate but related question concerns route choice. When drivers are not just following each other but also selecting paths through a network, congestion acquires a game-theoretic dimension. We formalize this through congestion games and Wardrop equilibria [2, 17, 23], and we simulate four routing strategies on a grid network to quantify how congestion externalities accumulate as demand grows. A central finding is a demand threshold beyond which Braess's paradox [3] emerges. Adding a road to the network strictly worsens everyone's travel time because selfish drivers collectively overload the new shortcut. This illustrates that congestion is not only a car-following instability but also a coordination failure, one that persists even when the road infrastructure improves.

To bridge the gap between these discrete agent interactions and continuous network-wide dynamics, we introduce a macroscopic optimal-transport framework. While microscopic and game-theoretic models address the mechanistic how of congestion, this top-down layer explains the organizing variational principles. Grounded in Mean Field Game theory [12] and Wasserstein gradient flows [1, 8], this multi-commodity model represents heterogeneous vehicle populations through a global free energy functional. By balancing entropic diffusion, destination potentials, and nonlocal congestion penalties, this continuum approach recovers the empirical fundamental diagrams and captures emergent spatial properties, such as capacity hysteresis and bidirectional phase-locking. At the same time, it naturally avoids the pathological coordination failures seen in selfish routing.

The rest of the paper is organized as follows. Section 2 reviews related work on microscopic traffic models, reinforcement learning for traffic control, and game-theoretic routing. Section 3 presents the three microscopic models and the unified simulation interface. Section 4 reports simulation results and parameter analyses for each model. Section 5 develops the reinforcement

learning approach and presents evaluation results. Section 7 covers the game-theoretic routing framework and strategy comparisons. Section 8 formally details the macroscopic optimal-transport model and multi-commodity network evaluations. Section 9 discusses implications across all modeling scales, and Section 10 concludes.

2. Related Work

Microscopic traffic models. Microscopic traffic models describe individual vehicles and their interactions with nearby traffic [7]. Cellular automaton models, especially the Nagel–Schreckenberg model, discretize both space and time and update vehicles through acceleration, braking, randomization, and movement rules [15]. Their simplicity makes them effective for studying exclusion effects, stochastic braking, and phase transitions. Continuous or semi-continuous car-following models instead describe vehicle motion through velocities, accelerations, headways, and safety constraints. The Krauss model computes a stochastic safe speed that preserves collision-free dynamics [10], while the Intelligent Driver Model defines acceleration as a nonlinear function of speed, gap, and relative velocity [20]. These models differ in their microscopic assumptions but can produce qualitatively similar macroscopic regimes.

Fundamental diagrams, phase transitions, and jam morphology. The fundamental diagram, relating density to flow, is a central object in traffic-flow theory [5, 7]. It summarizes the transition from low-density free flow to high-density congestion and is closely connected to classical kinematic-wave theory [13, 16]. However, one-dimensional flow-density curves can hide important dependencies on driver behavior, stochasticity, safety parameters, and local heterogeneity. Empirical and theoretical work on stop-and-go waves and congested phases shows that traffic breakdown is also a spatiotemporal phenomenon [9, 19]. For this reason, our benchmark complements fundamental diagrams with phase diagrams, slow-car maps, jam detectors, velocity variance, local density dispersion, and safety-critical observables such as time-to-collision [6]. Density-based clustering and connected-component methods are used to summarize jam morphology, with DBSCAN providing a generic reference for density-based cluster extraction [4].

Reinforcement learning and mixed-autonomy traffic control. A growing literature studies whether a small fraction of autonomous or controlled vehicles can stabilize traffic and improve aggregate flow [11, 22, 24]. Reinforcement learning is natural in this setting because a controlled vehicle acts under partial observations while optimizing delayed traffic-level rewards [14]. Mixed-autonomy studies show that learned or optimized driving policies can dissipate stop-and-go waves, smooth velocity profiles, and improve throughput even when most vehicles remain uncontrolled [11, 22]. Our work follows this line but embeds the controller in the same benchmark used for classical traffic models, allowing the control results to be interpreted relative to the underlying microscopic congestion mechanisms.

Congestion games and selfish routing. Traffic congestion also arises from route-choice interactions. Congestion games model drivers as self-interested agents whose costs depend on the number of users sharing each road segment [17]. In large populations, Wardrop equilibria describe flows in which all used routes have minimal travel time [23]. Under monotone edge costs, the Beckmann transformation expresses the equilibrium as the solution of a convex optimization problem [2]. Nevertheless, selfish equilibria generally differ from the social optimum, and the Price of Anarchy quantifies the resulting inefficiency [18]. Braess's paradox provides a classical example in which adding a road can worsen equilibrium travel time [3]. This literature complements microscopic car-following models by explaining network-level congestion as a strategic coordination failure.

Macroscopic traffic flow, optimal transport, and mean field games. Macroscopic traffic models describe vehicle populations through density and flow fields rather than individual trajectories. Classical first-order models such as Lighthill–Whitham–Richards represent traffic through conservation laws and density-dependent velocity relations [13, 16]. Our graph-density model follows a different but related variational viewpoint: traffic density evolves according to a free-energy functional containing entropy, destination attraction, congestion interaction, and capacity pressure. This connects to Wasserstein gradient flows and the Jordan–Kinderlehrer–Otto variational time-discretization [1, 8]. In the multi-commodity setting, where several origin–destination groups interact through shared capacity, the model also relates to mean field game theory [12]. This perspective provides a bridge between individual route incentives, collective density evolution, and network-level congestion pressure.

3. Problem Setting and Traffic Models

3.1. General traffic simulation setting

We consider a set of vehicles moving on a road network. Each vehicle i has a position $x_i(t)$, a velocity $v_i(t)$, and possibly additional internal state variables depending on the model. The road geometry may be a ring road, a bottleneck, a multilane road, or a graph extracted from a road network.

The goal is to compare how different microscopic update rules generate macroscopic traffic behavior under shared initial conditions, densities, perturbations, and measurement protocols, following the standard microscopic-traffic perspective reviewed in [7]. We denote by $M \in \mathcal{M}$ a microscopic traffic model, where

$$\mathcal{M} = \{\text{Nagel–Schreckenberg}, \text{Krauss}, \text{Intelligent Driver Model}\}.$$

3.2. Nagel–Schreckenberg Cellular Automaton

The Nagel–Schreckenberg model discretizes a road as a one-dimensional lattice of L cells with periodic boundary conditions [15]. Each vehicle i occupies a single cell $x_i(t)$ and carries an integer velocity $v_i(t) \in \{0, \dots, v_{\max}\}$. The gap $g_i(t)$ is defined as the number of empty cells between vehicle i and the vehicle immediately ahead.

At each time step, all vehicles update synchronously through four rules:

1. **Acceleration:**
$$v_i \leftarrow \min(v_i + 1,\, v_{\max})$$
2. **Braking:**
$$v_i \leftarrow \min(v_i,\, g_i)$$
3. **Randomization:** with probability p_{slow}, if $v_i > 0$:
$$v_i \leftarrow v_i - 1$$
4. **Movement:**
$$x_i \leftarrow x_i + v_i \pmod{L}$$

The randomization step is the model's core mechanism: by stochastically reducing a vehicle's speed, it breaks deterministic flow and produces spontaneous jam formation, that is, stop-and-go waves that emerge from purely local interactions without any global bottleneck. Despite its simplicity, Nagel–Schreckenberg reproduces characteristic traffic regimes and serves as the computational baseline for the models introduced below [7, 15].

3.3. Krauss stochastic safe-speed model

The Krauss model is a microscopic car-following model that computes a stochastically perturbed velocity while enforcing collision avoidance through a linearized kinematic constraint [10]. At each discrete time step, the state of vehicle i is updated based on its current velocity $v_i(t)$, maximum acceleration a, reaction time τ, and the gap $g_i(t)$ to the leading vehicle $i+1$ with velocity $v_{i+1}(t)$. The safe speed is computed via the following linear approximation, which assumes both vehicles decelerate symmetrically over $\tau + 1$ steps [10]:

$$v_{\text{safe},i}(t) = \max\left(0,\ v_{i+1}(t) + \frac{g_i(t) - v_{i+1}(t)\,\tau}{\tau + 1}\right).$$

The desired speed is then bounded by the safe speed, the road speed limit $v_{\max}$, and the vehicle's acceleration capability:

$$v_{\text{des},i}(t) = \min(v_i(t) + a,\ v_{\text{safe},i}(t),\ v_{\max}).$$

A stochastic perturbation models driver imperfection. The noise term $\eta_i(t) \sim \mathcal{U}(0, \varepsilon)$, where $\varepsilon \in [0, v_{\max}]$ scales the amplitude of behavioral noise:

$$v_i(t+1) = \max(0,\ v_{\text{des},i}(t) - \eta_i(t)).$$

Finally, position is updated by:

$$x_i(t+1) = x_i(t) + v_i(t+1) \pmod{L}.$$

The model parameters are summarized in Table 1.

Parameter	Description	Default value
$v_{\max}$	Maximum speed	5 cells/step
a	Maximum acceleration	1 cell/step^2
τ	Reaction time	0.5 steps
ε	Noise amplitude	0.3 cells/step

Table 1: Krauss model parameters used in simulations.

3.4. Intelligent Driver Model continuous car-following model

The Intelligent Driver Model [20] is a continuous car-following model in which the acceleration of vehicle i depends on its current velocity v_i, the gap s_i to the leading vehicle, and the approach rate $\Delta v_i = v_i - v_{i+1}$. The acceleration law is:

$$\dot{v}_i = a\left[1 - \left(\frac{v_i}{v_0}\right)^{\delta} - \left(\frac{s^*(v_i, \Delta v_i)}{s_i}\right)^2\right],$$

where the desired gap s^* is:

$$s^*(v_i, \Delta v_i) = s_0 + v_i T + \frac{v_i\,\Delta v_i}{2\sqrt{ab}}.$$

Here v_0 is the desired free-road speed, a is the maximum acceleration, b is the comfortable deceleration, T is the desired time headway, s_0 is the minimum spacing at standstill, following the standard IDM parametrization [20]. (set to one cell length in the discrete implementation), and $\delta = 4$ controls the sharpness of the free-road acceleration profile.

The continuous dynamics are integrated using an explicit Euler scheme with time step $\Delta t = 1$ step:

$$v_i(t+1) = \text{clip}(v_i(t) + \dot{v}_i(t),\ 0,\ v_0)\,, \qquad x_i(t+1) = x_i(t) + v_i(t+1) \pmod{L}.$$

A hard safety constraint is applied after integration to prevent rear-end collisions:

$$v_i(t+1) \leftarrow \min(v_i(t+1),\ s_i(t))\,.$$

An optional Gaussian noise term $\eta_i(t) \sim \mathcal{N}(0, \varepsilon^2)$ can be added to $\dot{v}_i$ to model stochastic driver behavior. The model parameters are summarized in Table 2.

Parameter	Description	Default value
v_0	Desired free-road speed	5.0 cells/step
a	Maximum acceleration	1.5 cells/step^2
b	Comfortable deceleration	3.0 cells/step^2
s_0	Minimum gap at standstill	1.0 cell
T	Desired time headway	0.5 steps
δ	Acceleration exponent	4
ε	Noise amplitude	0.1 cells/step^2

Table 2: IDM parameters used in simulations.

3.5. Theoretical Properties of Evaluated Models

To provide a baseline for our empirical observations, we briefly summarize the theoretical steady-state properties of the three traffic models implemented in our benchmark.

Nagel-Schreckenberg (NaSch) Model The NaSch model [15] is a stochastic cellular auto maton known for reproducing the phase transition between free-flow and congested traffic. In the purely deterministic limit (randomization parameter $p = 0$), the system exhibits a sharp phase transition at a critical density ρ_c. Below ρ_c, all vehicles travel at the maximum allowed velocity $v_{\max}$ without interacting. Above this threshold, congestion inevitably forms. The critical density is theoretically defined as $\rho_c = \frac{1}{v_{\max}+1}$ (see Appendix A.1). Introducing stochastic braking ($p > 0$) removes this sharp transition, leading to spontaneous jam formations even below ρ_c and reducing the overall capacity of the road.

Krauss Model The Krauss model [**?**] bridges the gap between cellular automata and continuous car-following models by calculating a bounded safe velocity at each discrete time step. The theoretical foundation of the model rests on the guarantee of collision-free dynamics, provided that vehicles do not exceed a maximum deceleration b. The equilibrium state of the Krauss model dictates that vehicles maintain a safe gap proportional to their velocity and the reaction time of the driver (τ). This mechanical constraint inherently bounds the maximum theoretical throughput of the road segment.

Intelligent Driver Model (IDM) Unlike the discrete rule sets of NaSch and Krauss, the IDM [21] is a continuous-time, continuous-space car-following model. It is characterized by its smooth transition from acceleration in free flow to deceleration upon approaching a slower vehicle. The theoretical equilibrium state of the IDM (where all vehicles travel at a constant velocity v_e with zero acceleration and zero speed difference) yields a deterministic fundamental diagram. This allows for an exact analytical calculation of the equilibrium gap $s_e(v)$ and the resulting macroscopic density ρ (see Appendix A.2).

4. Simulation and Results

4.1. Simulation Setup and Protocol

The primary geometry used for the benchmark simulations is a single-lane ring road, a standard closed topology for studying self-organized traffic jams without external bottlenecks [7, 19]. This closed-topology approach is ideal for isolating the spontaneous emergence of "phantom traffic jams" and studying system stability without the compounding variables of open boundaries or inflow/outflow fluctuations.

The road is spatially discretized into 200 individual cells, where each cell can be occupied by at most one vehicle to enforce the exclusion principle, following the cellular-automaton traffic modeling framework introduced by Nagel and Schreckenberg [15]. Periodic boundary conditions are applied, meaning a vehicle exiting the final cell immediately re-enters the first, ensuring that the global density ($\rho = N/L$) remains strictly constant throughout the duration of each simulation run.

The common benchmark protocol executed for these simulations consists of the following sequence:

1. Selecting a traffic model $M \in \mathcal{M}$;
2. Selecting a road geometry and setting the initial density;
3. Running the model for a warm-up period to reach steady-state dynamics;
4. Collecting trajectories and observables during the measurement period;
5. Repeating the process over varying random seeds and parameter grids;
6. Comparing aggregate metrics across models and scenarios.

4.2. Observables and Metrics

To evaluate traffic dynamics and model behavior, we utilize a suite of metrics that capture both the macroscopic traffic state of the system and the microscopic safety of individual driving policies [7].

- **Macroscopic Traffic State:** We track fundamental global quantities including global density (ρ), mean velocity ($\langle v \rangle$), and aggregate flow (J), which are the classical variables underlying empirical and theoretical fundamental diagrams [5, 7]. Additionally, we monitor velocity variance and local density heterogeneity to identify flow instability and the transition from free-flowing traffic to synchronized or congested regimes [9].
- **Jam Dynamics and Propagation:** Jams are defined as persistent spatiotemporal structures rather than transient slowdowns, consistent with empirical observations of stop-and-go waves and traffic breakdown [9, 19]. By applying density and flux thresholds over sliding spatial windows, we isolate congested clusters. This allows us to extract morphological statistics such as the total jam count, mean jam size, and the characteristic backward wave speed (c) at which congestion propagates through the system.
- **Safety and Risk Assessment:** The safety profile of the simulated driving behavior is quantified through three operational risk metrics:
  - *Collision Rates:* The frequency of events where the spatial gap between consecutive vehicles drops to zero.
  - *Emergency Braking:* The occurrence of sudden, high-magnitude velocity drops within a single time step, indicating poor anticipation or local flow instability.

– *Time-to-Collision (TTC):* A continuous surrogate safety measure used to capture "near-miss" scenarios by evaluating closing speeds relative to spatial gaps, identifying high-risk interactions before collisions occur [6].

Algorithm 1 Generic traffic benchmark protocol

```
Select traffic model M, geometry G, density ρ, and parameter vector θ
Initialize N = ρL vehicles on G
for t = 1 to T_warmup do
    Update all vehicles according to model M
end for
for t = 1 to T_eval do
    Update all vehicles according to model M
    Record positions, velocities, flow, and jam indicators
end for
Aggregate metrics over time and random seeds
return mean speed, flow, jam statistics, spacetime trajectories, and safety metrics
```

4.3. Macroscopic Regime Classification

Using the macroscopic observables, we classify traffic into three qualitative regimes based on the joint analysis of the mean velocity $\langle v \rangle$ and velocity variance $\mathrm{Var}(v)$, following the common distinction between free-flow, metastable stop-and-go, and congested traffic states [7, 9]:

- **Free flow:** $\langle v \rangle \approx v_{\max}$ and $\mathrm{Var}(v) \approx 0$.
- **Metastable / Stop-and-go:** Characterized by high $\mathrm{Var}(v)$ with the spontaneous formation of localized, high-density clusters.
- **Congested flow:** $\langle v \rangle \to 0$ at high density, characterized by a global collapse of throughput.

4.4. Fundamental Diagram Estimation

The fundamental diagram describes the steady-state relationship $\rho \mapsto J(\rho)$ and is one of the classical empirical summaries of road traffic flow [5, 7]. To systematically construct this empirical relation across models, our estimation protocol consists of the following parameters:

1. **Density Sweep:** The global density is varied within $\rho \in [0.05, 0.95]$ using a step size of 0.05.
2. **Measurement:** For each density ρ, we run $M = 20$ independent simulations using different random seeds to account for stochastic variations.
3. **Stationarity:** We discard the first $T_{\text{warmup}} = 10^3$ simulation steps to remove transient dynamics before measuring the steady-state regime.
4. **Averaging:** Flow and velocity metrics are averaged over a measurement horizon of $T_{\text{horizon}} = 5 \cdot 10^3$ steps.

4.5. Nagel–Schreckenberg Model Simulation Results

The Nagel–Schreckenberg model represents a discrete cellular-automaton approach in which traffic dynamics emerge from local acceleration, braking, randomization, and movement rules [15]. Its behavior is dictated by the balance between deterministic flow and stochastic disruptions.

4.5.1. Macroscopic Regime and Phase Transition

To understand the global impact of stochastic noise, we map the system's macroscopic states across varying densities (ρ) and braking probabilities (p_{slow}), extending the classical density-flow analysis of the Nagel–Schreckenberg model [7, 15]. Figure 1 presents the resulting phase diagram, effectively capturing the transition from stable to congested flow without the need for isolated metric tracking.

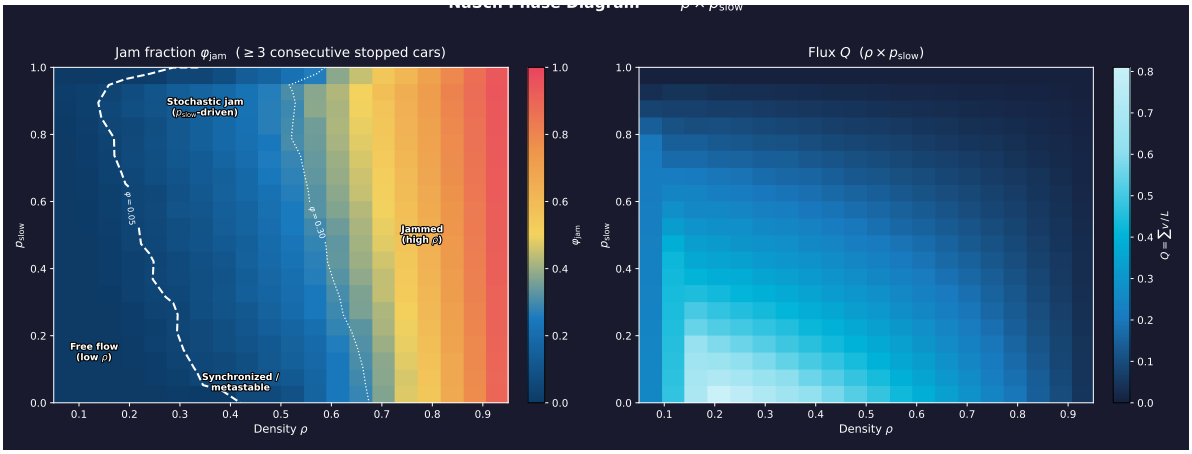


Figure 1: **NaSch phase diagram in** (ρ, p_{slow}) **space.** The free-flow region (low density, low noise) is separated from the jammed region by a boundary that shifts to lower density as p_{slow} increases, demonstrating how stochastic noise promotes early jam nucleation. The heatmaps reflect the macroscopic flux collapse and the concurrent rise in jam density (ϕ_{jam}).

The phase diagram illustrates a clear boundary between laminar motion and breakdown, which can be categorized into four distinct macroscopic regimes. At low density and low noise, the system remains in a **free flow** state, where vehicles maintain near-maximum velocity and throughput scales linearly. As density increases under very minimal noise ($p_{\text{slow}} \approx 0$), the flow enters a **synchronized/metastable** state; throughput remains high, but the system becomes highly sensitive to perturbations.

Crucially, as p_{slow} increases, the critical density drops significantly, pushing the system into a **stochastic jam (p_{slow}-driven)** regime. Here, high randomization causes spontaneous decelerations that nucleate phantom jams even at moderate densities. Finally, across all noise levels at high densities ($\rho \gtrsim 0.7$), the system enters a fully **jammed** state. The right panel's flux heatmap shows the collapse of global throughput in these regions, corresponding perfectly with the left panel's surge in ϕ_{jam} (defined as the percentage of localized clusters with at least three consecutive stopped vehicles). This joint visualization confirms that high stochastic noise severely limits capacity by prematurely fragmenting the flow into stop-and-go waves.

4.5.2. Spatiotemporal Dynamics and Jam Morphology

While the phase diagram defines the macroscopic boundaries of congestion, the spatiotemporal trajectories reveal the microscopic mechanics of these jams. We examine the morphological impact of both the stochastic noise (p_{slow}) and the physical speed limit ($v_{\max}$).

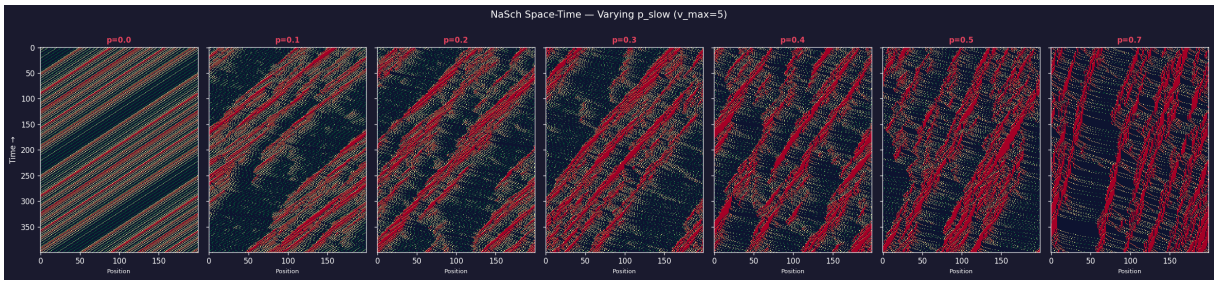


Figure 2: **NaSch Space-Time (p_{slow})**: Evolution of trajectories for varying p_{slow}. High randomization values disrupt synchronized flow, transforming laminar motion into a fragmented pattern of stop-and-go waves propagating upstream.

As illustrated in Figure 2, increasing p_{slow} directly drives the nucleation of phantom jams, consistent with the role of random braking in the original Nagel–Schreckenberg model [15]. At low noise levels, perturbations are minor and quickly dissipate. Above the critical threshold mapped in the phase diagram, the system undergoes a sharp morphological transition: spontaneous decelerations compound, creating the characteristic backward-propagating sawtooth patterns inherent to congested discrete systems. The spatial homogeneity is entirely lost to a highly fragmented stop-and-go regime.

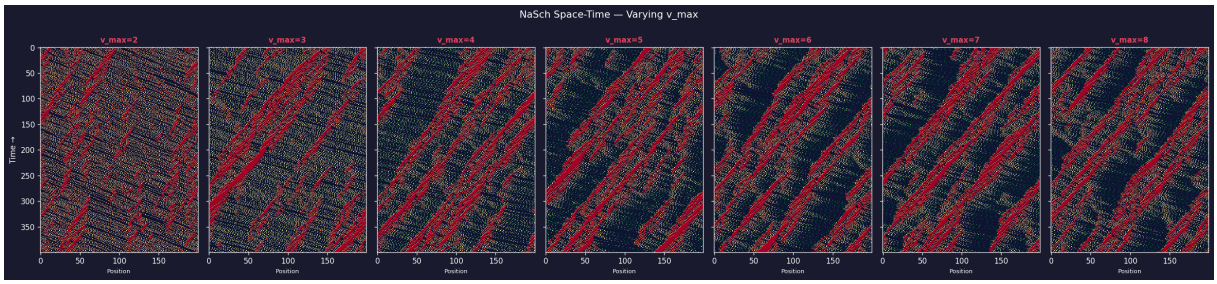


Figure 3: **NaSch Space-Time ($v_{\max}$)**: Trajectories for varying speed limits. Higher $v_{\max}$ increases the void space between vehicles required for safety, steepening free-flow slopes while amplifying jam severity.

Conversely, altering the maximum velocity $v_{\max}$ shifts the geometric scale of these interactions, as seen in Figure 3. Higher $v_{\max}$ values yield steeper trajectory slopes during free-flow phases, reflecting faster vehicle propagation and higher theoretical throughput. However, these higher speeds inherently increase the required safety gaps. Consequently, the system becomes highly sensitive to perturbations: when a fast-moving vehicle encounters a slow cluster, the speed differential is immense. This produces much sharper deceleration fronts, severely reducing the Time-to-Collision (TTC) and generating denser, more severe backward-propagating waves compared to lower speed limits.

4.5.3. Detailed Safety and Variance Metrics

While the phase diagram and spatiotemporal plots capture the macroscopic flow and morphological shape of congestion, the underlying safety profile of the NaSch model is best understood through its microscopic operational metrics (detailed comprehensively in Appendix C).

As stochastic noise (p_{slow}) increases, we observe a sharp peak in velocity standard deviation precisely at the phase transition boundary, accompanied by a surge in local density dispersion. Safety metrics deteriorate rapidly in this regime: gap-zero events and emergency braking maneuvers rise sharply, while the average Time-to-Collision (TTC) plummets, reflecting the severity of the sudden, uncoordinated decelerations inherent to high-noise cellular automata.

Similarly, scaling the maximum velocity ($v_{\max}$) exposes a safety trade-off. While higher speed limits increase the theoretical free-flow capacity, they drastically amplify the severity of interactions at jam boundaries. The larger speed differentials force sharper braking events, leading to a higher frequency of near-critical (low TTC) situations despite the larger spatial gaps maintained during free flow.

4.6. Krauss Model Simulation Results

The Krauss model introduces continuous dynamics while maintaining a strict linearized safety constraint to prevent collisions. Its behavior is dictated by the interplay between stochastic perturbations (noise) and the drivers' reaction times, governing the balance between flow capacity and driving caution.

4.6.1. Macroscopic Regime and Phase Transition

To evaluate the global impact of stochastic noise (ϵ), we map the system's macroscopic states across varying densities and noise levels. Figure 4 presents the Krauss phase diagram, capturing the transition from quasi-deterministic car-following to a stochastically congested regime.

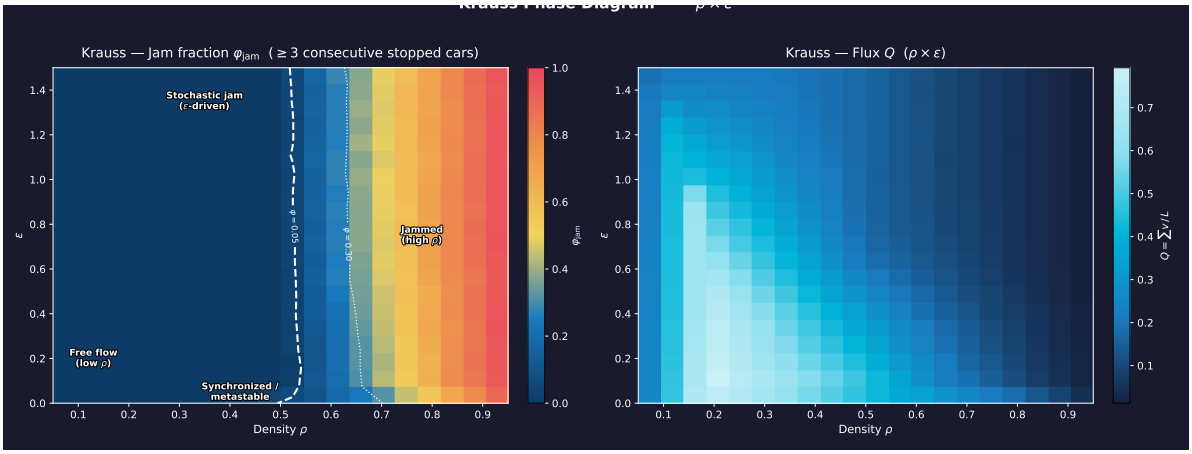


Figure 4: **Krauss phase diagram.** The kinematic safety constraint softens the phase boundary relative to NaSch: transitions are more gradual and the fully jammed region is smaller at equivalent noise levels.

Unlike the sharp breakdown observed in the discrete NaSch model, the Krauss model exhibits a much softer phase boundary. The continuous acceleration profile and kinematic safety constraint ensure that velocity transitions are less abrupt. While throughput collapses at high density and high noise, the transition is gradual, and the fully jammed parameter space is notably smaller than in purely discrete cellular automata.

4.6.2. Spatiotemporal Dynamics and Jam Morphology

The structural manifestation of these jams heavily depends on the noise amplitude (ϵ) and the prescribed reaction time (τ). We examine both dimensions through their spatiotemporal trajectories.

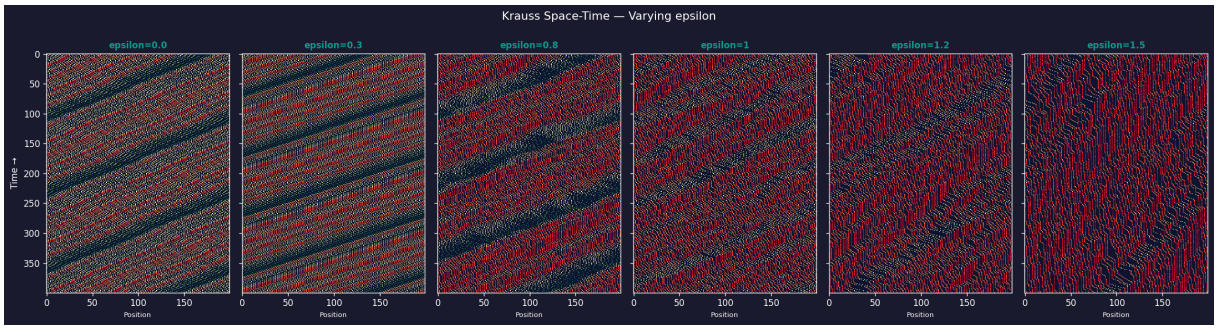


Figure 5: **Krauss Space-Time (ϵ)**: Higher noise levels compromise string stability, yet the safety-speed rule ensures collision-free motion is maintained. Jam boundaries are smoother than in NaSch due to continuous velocity updates.

Figure 5 illustrates the impact of stochastic noise. At low ϵ, the system converges toward a near-uniform, synchronized flow. As ϵ increases, micro-perturbations accumulate and compromise string stability, nucleating persistent stop-and-go waves. However, because of the continuous velocity updates and strict braking bounds, the backward-propagating shockwaves have significantly smoother boundaries and less severe internal compression than those in the NaSch model.

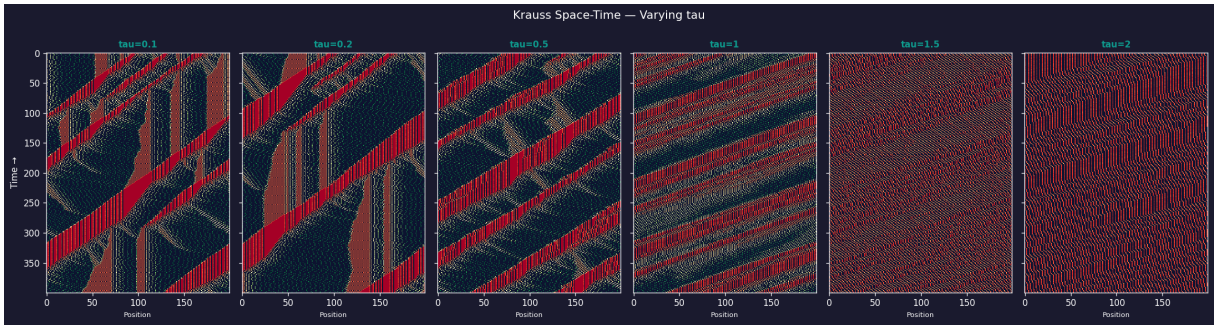


Figure 6: **Krauss Space-Time (τ)**: As reaction time increases, the system's ability to dissipate perturbations decreases. Large τ forces conservative headways, reducing capacity and producing wide, persistent congestion.

Figure 6 isolates the effect of the reaction time parameter (τ). Large τ values force the model's safe speed calculation (v_{safe}) to become highly conservative. Vehicles are required to maintain significant inter-vehicle distances, which fundamentally reduces road capacity. When perturbations occur, vehicles decelerate well in advance of the obstacle. This generates wide, slowly propagating jam fronts rather than sharp, localized nucleation events, resulting in extended regions of synchronized slow traffic.

4.6.3. Detailed Safety and Variance Metrics

The underlying microscopic safety profile of the Krauss model (detailed comprehensively in Appendix D) reveals an interesting trade-off between capacity and risk.

As noise (ϵ) increases, mean speed remains remarkably robust due to the safety constraint capping extreme decelerations, though velocity standard deviation grows. Safety metrics only degrade (showing increased emergency braking and reduced Time-to-Collision) beyond a critical noise amplitude.

Conversely, the reaction time (τ) presents a paradox. High τ values drastically reduce road capacity and plummet the mean velocity, as vehicles are constrained to lower speeds even in low-density conditions. However, because vehicles maintain massive safety margins, gap-zero

events and emergency braking are almost entirely eliminated, yielding a highly safe but deeply inefficient macroscopic flow.

4.7. Parameter Analysis: Intelligent Driver Model (IDM)

The Intelligent Driver Model (IDM) is a continuous car-following model whose rich parameter space controls the interplay between free-road acceleration, desired headway, and anticipatory braking. Unlike the NaSch and Krauss models, where congestion is primarily noise-driven, phantom jams in the IDM can emerge from purely deterministic dynamics through linear string instability.

Algorithm 2 Generic traffic benchmark protocol

Select traffic model M, geometry G, density ρ, and parameter vector θ
Initialize $N = \rho L$ vehicles on G
for $t = 1$ to T_{warmup} **do**
    Update all vehicles according to model M
end for
for $t = 1$ to T_{eval} **do**
    Update all vehicles according to model M
    Record positions, velocities, flow, and jam indicators
end for
Aggregate metrics over time and random seeds
return mean speed, flow, jam statistics, spacetime trajectories, and safety metrics

4.7.1. Macroscopic Regime and Deterministic Phase Transition

The most distinctive feature of the IDM is its deterministic phase transition, governed primarily by the desired time headway (T). Figure 7 maps this macroscopic behavior, illustrating how the boundary between stable uniform flow and unstable oscillatory motion shifts based on driver caution.

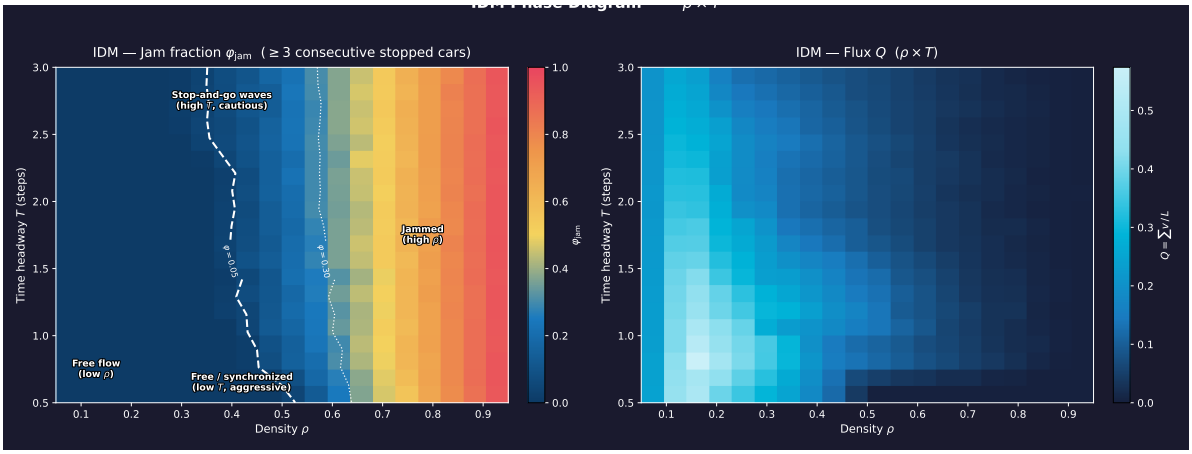


Figure 7: **IDM phase diagram in (ρ, T) space.** The dashed curves indicate the empirical transition thresholds ($\phi = 0.05$ and $\phi = 0.30$). Highly cautious driving (high T) demands larger spatial gaps, effectively reducing capacity and causing the system to break down into stop-and-go waves at much lower densities than aggressive parameterizations.

The phase diagram divides the parameter space into four identifiable regimes. At low densities, the system maintains a robust **free flow (low ρ)** state regardless of the time headway. As density increases to intermediate levels, the impact of T dictates the flow stability. With

low time headways ($T \lesssim 1.0$), the aggressive driving behavior efficiently utilizes physical space, allowing the system to maintain a stable **free / synchronized** flow where perturbations are safely absorbed.

Conversely, if the time headway is high ($T \gtrsim 1.5$), the highly cautious driving behavior demands larger safety gaps, severely reducing the road's effective capacity. When the actual density exceeds this reduced capacity, the system crosses the stability threshold. The nonlinear feedback of the desired gap calculation exponentially amplifies small perturbations, forcing the system to collapse into **stop-and-go waves (high T, cautious)**. Finally, as physical space is exhausted at extreme densities ($\rho \gtrsim 0.65$), the system collapses into a completely **jammed (high ρ)** state across all time headway parameterizations.

4.7.2. Spatiotemporal Dynamics and Jam Morphology

The rich parameter space of the IDM allows us to isolate how different driving behaviors physically shape congestion. We evaluate these morphological changes through the system's spatiotemporal trajectories.

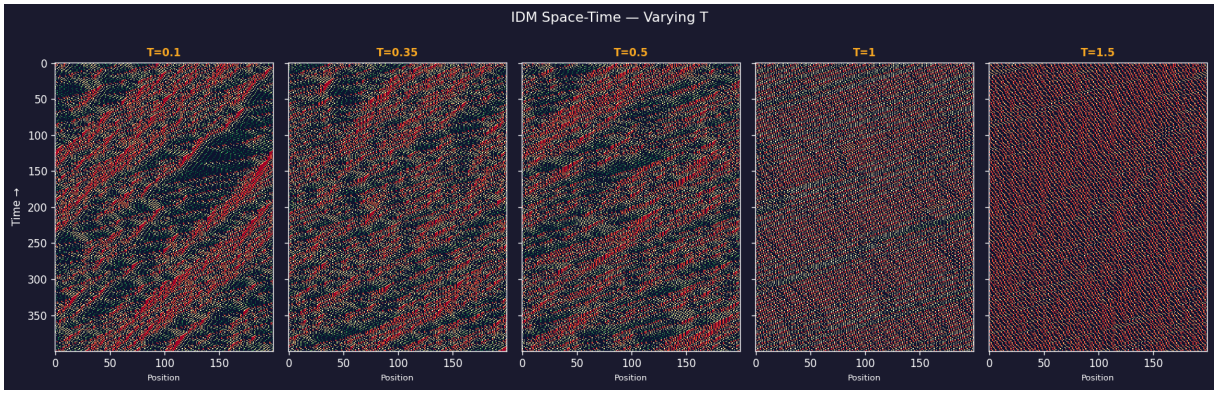


Figure 8: **IDM Space-Time (T)**: Transition from unstable oscillatory motion at low T to a stable car-following state at high T. The critical headway marks a genuine phase transition driven by linear instability.

Instability and Noise: Figure 8 visualizes the aforementioned T-driven phase transition. At critically low T, the system collapses into a single, massive jam. Interestingly, while the IDM can jam deterministically, Figure 9 demonstrates that adding stochastic noise (ϵ) acts as a rapid seed for nucleation. In near-critical regimes, even microscopic noise is sufficient to trigger the IDM's nonlinear amplification, producing spatially fragmented, macroscopic waves much earlier than purely deterministic conditions would dictate.

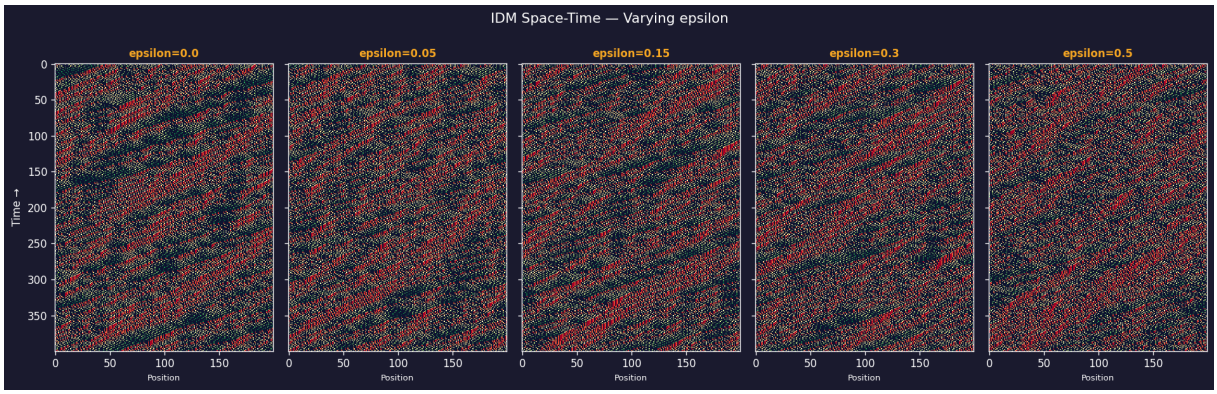


Figure 9: **IDM Space-Time (ϵ)**: Emergence of spontaneous jams from microscopic fluctuations. Even small noise is sufficient to seed the IDM's nonlinear amplification in near-critical parameter regimes.

Acceleration and Speed Trade-offs: Figure 10 and Figure 11 showcase the impact of maximum acceleration (a) and maximum speed ($v_{\max}$). High a values accelerate vehicle recovery after braking but increase the coupling strength between vehicles, producing much sharper and more abrupt jam boundaries. Similarly, higher $v_{\max}$ increases free-flow throughput but drastically amplifies the severity of decelerations. The larger speed differential between free-flow and congested states forces sharper acceleration fronts and pronounced backward-propagating waves.

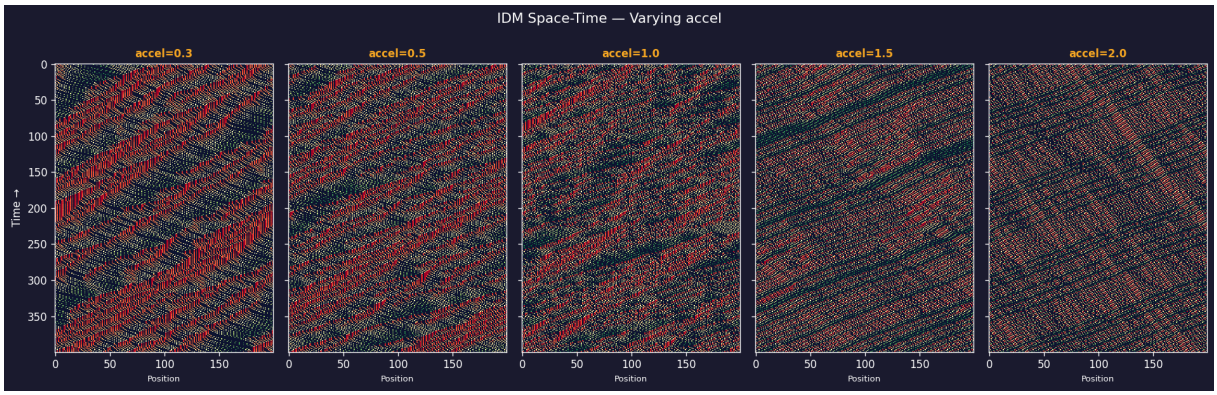


Figure 10: **IDM Space-Time (a)**: Effect of maximum acceleration. High a sharpens jam boundaries in the unstable regime while accelerating flow recovery in stable conditions.

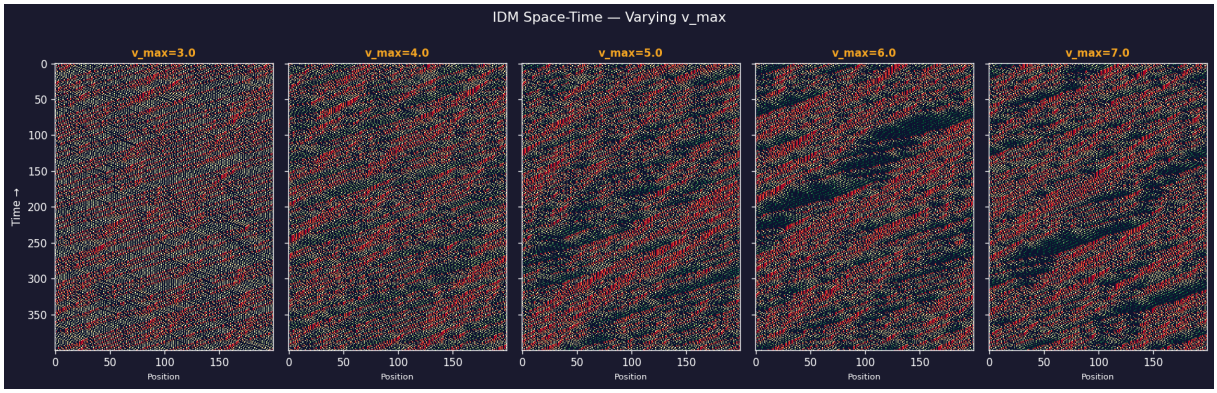


Figure 11: **IDM Space-Time ($v_{\max}$)**: Higher speed limits increase free-flow capacity but amplify the severity of deceleration at jam boundaries, producing sharper congestion fronts.

Physical Footprint: Finally, Figure 12 isolates the minimum spacing parameter (s_0). Larger s_0 forces vehicles to maintain greater gaps at a standstill. This reduces the maximum achievable density, shifting the fundamental diagram's congested branch and resulting in spatially extended, lighter jams rather than tightly packed clusters.

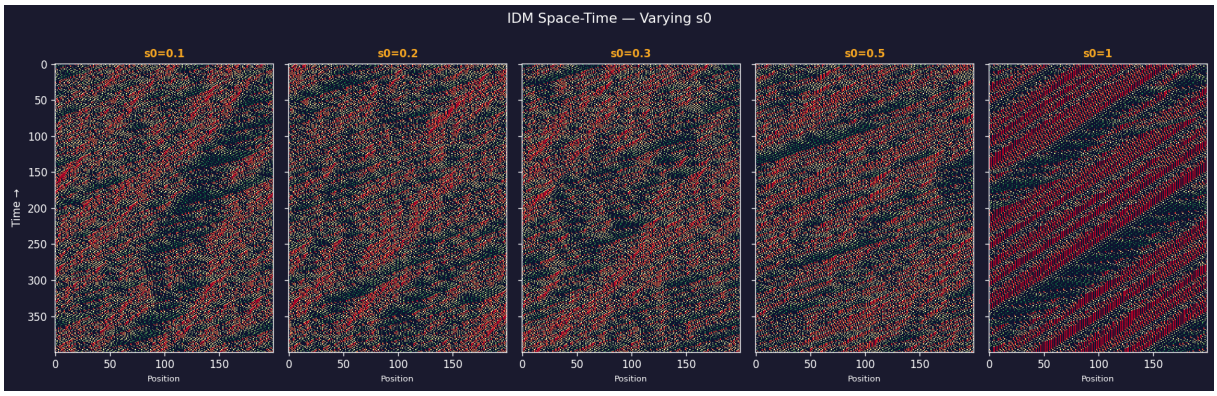


Figure 12: **IDM Space-Time (s_0)**: Effect of minimum spacing. Larger s_0 produces more spatially extended but less dense jams, reducing the maximum achievable congestion density.

4.7.3. Detailed Safety and Variance Metrics

The comprehensive microscopic metrics for the IDM (cataloged in Appendix E) reveal severe safety costs associated with aggressive parameter tuning.

In the unstable low-T regime, Time-to-Collision (TTC) degrades dramatically as vehicles are forced into frequent emergency braking events. While increasing maximum acceleration (a) or maximum speed ($v_{\max}$) improves mean velocity and recovery times, it explicitly trades away safety: gap-zero events surge as vehicles overshoot safe distances or fail to react to the high-speed differentials at jam boundaries. Conversely, increasing the minimum spacing (s_0) reliably improves safety margins and TTC, though at the direct cost of the road's congested capacity.

5. Reinforcement Learning Approach

5.1. Problem setting

$$\alpha = \frac{N_{\mathrm{RL}}}{N}.$$

The policy is trained with a single controlled vehicle ($N_{\mathrm{RL}} = 1$) and then deployed without retraining at larger values of α. The objective is to increase collective throughput and reduce speed variance, not to optimize the ego vehicle alone.

5.2. MDP formulation

Action space. Following the standard Markov decision process formulation used in reinforcement learning [14], all three models share the same three-action interface. Table 3 summarizes the mapping from discrete action to vehicle dynamics update.

Actions $a = 0$, $a = 1$, $a = 2$ correspond to deceleration, no change, and acceleration respectively. For Nagel–Schreckenberg, the residual correction is applied on top of the safe base velocity computed by the standard rule, so the learned action cannot cause a collision.

Observations. The default global observation vector gives the policy both ego-level and aggregate traffic information. Table 4 lists all features; the continuous models (Krauss and Intelligent

Table 3: Discrete action space and its effect for each traffic model.

Model	$\mathcal{A}$	Effect of action $a \in \{0,1,2\}$
Nagel–Schreckenberg	$\{0,1,2\}$	$v_{\text{RL}} = \text{clip}(v_{\text{base}} + (a-1),\ 0,\ v_{\max})$, gap constraint re-applied
Krauss	$\{0,1,2\}$	$v_{\text{des}} \leftarrow v_{\text{des}} + (a-1) \cdot 0.6$, then standard Krauss update
Intelligent Driver Model	$\{0,1,2\}$	$\dot{v} \leftarrow \dot{v} + (a-1) \cdot 0.35$, then standard Intelligent Driver Model integration

Driver Model) include two additional fields that have no direct equivalent in the cellular automaton.

Table 4: Global observation features. Checkmarks indicate inclusion; a dash indicates not applicable.

Feature	Nagel–Schreckenberg	Krauss	Intelligent Driver Model
Traffic density $\rho = N/L$	✓	✓	✓
Mean speed $\bar{v}/v_{\max}$	✓	✓	✓
Speed std $\sigma_v/v_{\max}$	✓	✓	✓
Flow $J/v_{\max}$	✓	✓	✓
Ego speed $v_{\text{ego}}/v_{\max}$	✓	✓	✓
Front gap $g_f/(L-1)$	✓	✓	✓
Rear gap $g_b/(L-1)$	✓	✓	✓
Fraction stopped f_{stop}	✓	✓	✓
Slowdown probability p_{slow}	✓	–	–
Speed scale $v_{\max}/7$	✓	✓	✓
Desired speed $v_0/v_{\max}$	–	✓	✓
Ego acceleration $\dot{v}$	–	–	✓
Total dimension	10	11	12

Reward. The reward is defined at the traffic level:

$$r_t = \text{clip}\left(2.6\frac{J_t}{v_{\max}} + 0.6\frac{\bar{v}_t}{v_{\max}} - 0.4\frac{\sigma_{v,t}}{v_{\max}} - 0.25\, f_{\text{stop},t} - 0.1\, b_t + 0.8\, \Delta J_t - 0.15\, |\Delta J_t|,\ -3,\ 3\right),$$

where $\Delta J_t = J_t - J_{t-1}$ is the flow increment and $b_t = \max(0,\ v_{\text{ego},t-1} - v_{\text{ego},t} - 0.5)$ penalizes hard braking. High throughput and smooth flow are rewarded; speed variance, stopped vehicles, and abrupt flow changes are penalized.

5.3. Training

The Q-network is a two-hidden-layer MLP (obs $\to$ 128 $\to$ 128 $\to$ 3, ReLU activations) trained with DQN [14]. The ϵ-greedy exploration rate is annealed from 1.0 to 0.05 over the first 60% of training steps, with a neutral-action bias during random exploration.

Curriculum. Rather than training on a fixed scenario, each episode samples a new traffic configuration: road length $L \in \{60, 80, 100, 120, 200\}$, density $\rho \sim \mathcal{U}(0.12, 0.90)$, slowdown probability $p_{\text{slow}} \sim \mathcal{U}(0.10, 0.45)$, and maximum speed $v_{\max} \in \{4, 5, 6, 7\}$. This forces the policy to generalize across traffic regimes rather than overfit to a single density or road geometry.

Algorithm 3 DQN training for one controlled vehicle (NaSch setting)

Initialize Q-network Q_ϕ, target network Q_{ϕ^-}, replay buffer $\mathcal{D}$
for episode $e = 1$ to E **do**
  Sample scenario: $(L, \rho, v_{\max}, p_{\text{slow}})$
  Initialize ring road; select one vehicle as controlled
  Run warm-up for T_{warmup} steps without learning
  for $t = 1$ to T **do**
    Observe o_t; select a_t via ϵ-greedy policy $Q_\phi(o_t, \cdot)$
    Apply residual action, enforce gap safety, advance all vehicles
    Compute r_t; store (o_t, a_t, r_t, o_{t+1}) in $\mathcal{D}$
    Sample mini-batch from $\mathcal{D}$; update Q_ϕ; periodically sync Q_{ϕ^-}
  end for
end for
return Q_ϕ

Hyperparameters. Key training settings: 600 episodes, 250 steps per episode, 40 warm-up steps per episode, replay buffer of 150 000 transitions, mini-batch size 256, discount $\gamma = 0.99$, Adam optimizer with learning rate 3×10^{-4}, Huber loss, target network updated every 500 steps, gradient clipping with max norm 5.

Evaluation metric. Performance is measured as the relative flow improvement over a matched all-baseline rollout at the same scenario and seed:

$$G = 100 \, \frac{J_{\text{RL}} - J_{\text{baseline}}}{|J_{\text{baseline}}|}.$$

6. RL Results

6.1. Effect of RL fraction on the fundamental diagram

Figures 13 14shows how the density-flow relationship changes as α increases from 0 to 1 for each model. In all three cases the mean flow rises with α across most densities.

The gain is most pronounced in Krauss, where the flow peak roughly triples at full deployment. For Intelligent Driver Model the improvement is smaller because the deterministic car-following already yields stable flow at moderate densities. Nagel–Schreckenberg shows intermediate behavior driven by the random slowdown mechanism.

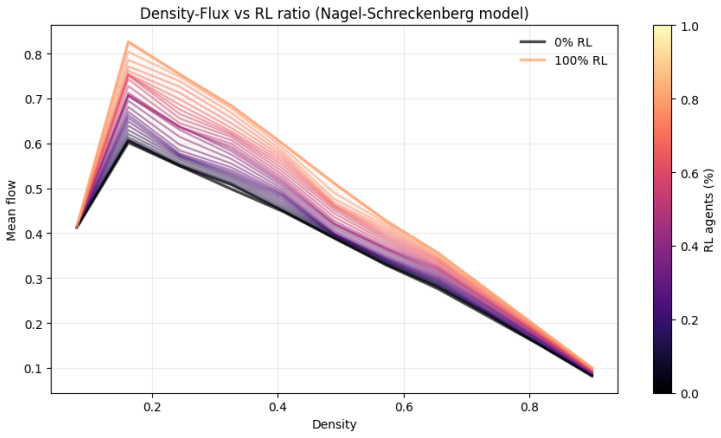
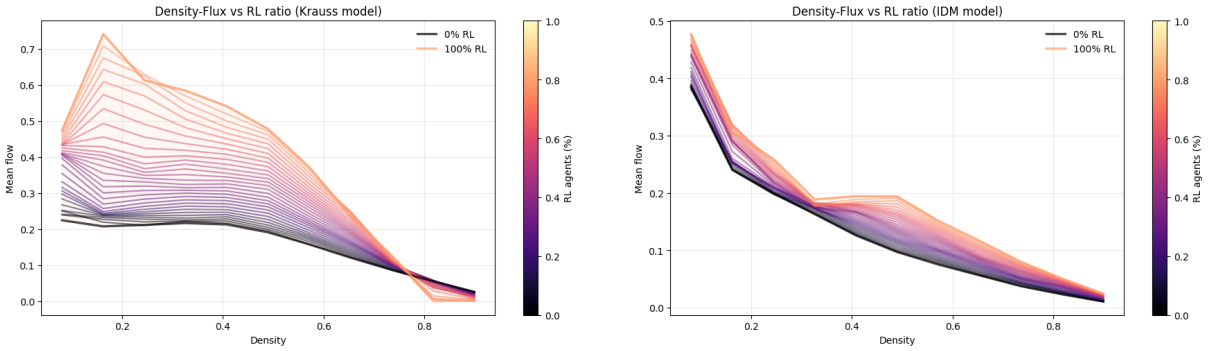


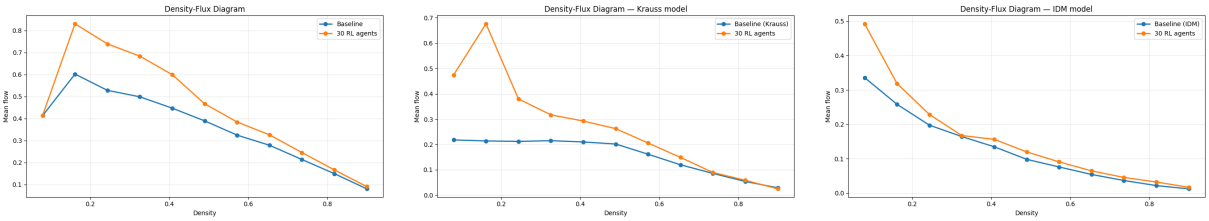
Figure 13: Nagel–Schreckenberg: α from 0 (black) to 1 (orange).



(a) Krauss: α from 0 to 1.

(b) Intelligent Driver Model: α from 0 to 1.

Figure 14: Density-flow diagrams for Krauss and IDM models as a function of RL fraction α.



(a) Nagel–Schreckenberg baseline vs. 30 RL agents.

(b) Krauss baseline vs. 30 RL agents.

(c) Intelligent Driver Model baseline vs. 30 RL agents.

Figure 15: Fundamental diagrams comparing the uncontrolled baseline (blue) to 30 RL-controlled vehicles (orange) for each traffic model.

6.2. Comparison at 30 RL agents

Figure 15 shows the density-flow curve for baseline traffic and for a deployment of 30 RL agents side by side for each model. The relative flow improvements at the evaluation density ($\rho \approx 0.30$) are reported in the spacetime figures below.

6.3. Space-time diagrams

Table 5 summarizes the relative flow gains at $\rho = 0.30$ for one and 30 RL agents per model. Figure 16 shows the corresponding space-time diagrams.

Table 5: Relative flow gain G at $\rho = 0.30$ for each model and RL deployment size.

Model	0 RL (baseline)	1 RL agent	30 RL agents
Nagel–Schreckenberg	0%	+5.4%	+43.1%
Krauss	0%	+16.5%	+117.7%
Intelligent Driver Model	0%	+5.5%	+32.7%

In the Nagel–Schreckenberg baseline, backward-propagating stop-and-go waves are clearly visible, consistent with classical observations of phantom jams on ring roads [19]. A single RL agent partially disrupts the jam structure (+5.4%); with 30 agents the waves are largely suppressed (+43.1%).

Krauss shows the largest absolute gain: the baseline produces persistent dense jams that 30 RL agents nearly eliminate (+117.7%). For Intelligent Driver Model, the deterministic car-following already yields a more regular wave pattern, so gains are more moderate (+32.7% with 30 agents).

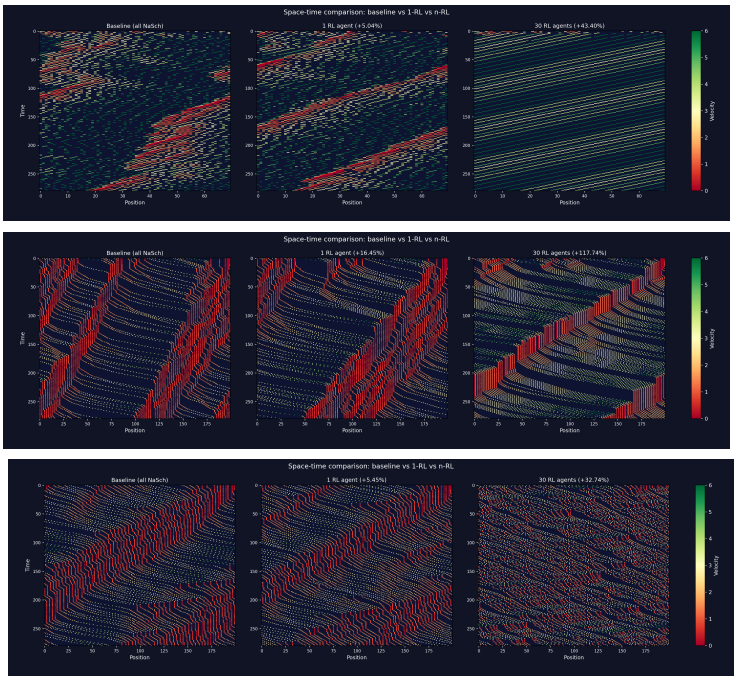


Figure 16: Space-time diagrams for Nagel–Schreckenberg (top), Krauss (middle), and Intelligent Driver Model (bottom). Each row shows baseline (left), one RL agent (center), and 30 RL agents (right). Vehicle color encodes speed from 0 (red) to $v_{\max}$ (green). Flow improvement percentages are shown in panel titles.

6.4. Observation-context ablation

A fundamental design choice in any partially observable traffic control problem is how much of the road state each agent can see [22, 24]. We ablate the observation radius r: setting $r = -1$ grants the agent access to *global* aggregate statistics (fleet-wide mean speed and mean gap), while $r \geq 1$ restricts each agent to the r nearest vehicles ahead and behind, yielding a fully local, scalable observation. The global setting provides an upper bound on information but cannot be deployed in a real system; the local settings are operationally realistic. All settings are evaluated over 50 random seeds to disentangle learning variance from structural differences.

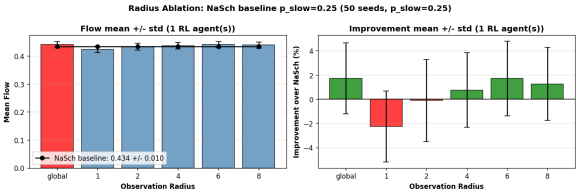
Figure 17 reports mean flow and relative improvement over the uncontrolled baseline for two noise levels ($p_{\text{slow}} \in \{0, 0.25\}$) and two fleet penetration rates (1 and 30 controlled vehicles out of 50).

Single-vehicle control. With a single controlled agent the global observation ($r = -1$) provides a clear advantage: the agent can track system-wide congestion and time its acceleration to absorb incoming perturbations before they amplify. Local radii $r \leq 3$ recover only a fraction of this gain, as the agent cannot anticipate disturbances that originate far upstream. Performance improves monotonically with r and plateaus around $r = 5$–7, suggesting that beyond a handful of neighbors the marginal value of additional context is small.

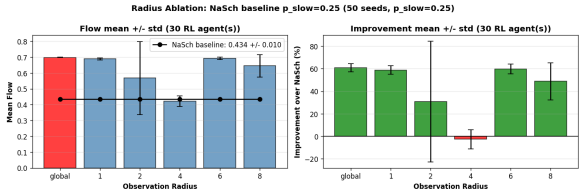
Multi-vehicle control. At 30 controlled vehicles the picture changes substantially. Because agents collectively span the ring, each one's local neighborhood is statistically representative of the global state, and the advantage of $r = -1$ over moderate local radii narrows considerably. At $p_{\text{slow}} = 0$ (no exogenous noise), a radius as small as $r = 2$ is sufficient to match global performance, confirming that in a deterministic environment the relevant signal is local. Under $p_{\text{slow}} = 0.25$, stochastic perturbations create transient heterogeneity that benefits from a slightly wider context ($r \approx 5$), but the gap relative to global observation remains modest.

Continuous-space robustness. Figure 18 replicates the radius sweep in the continuous-space environment. The qualitative ordering is preserved: global observation is best for sparse control, local radii suffice for dense fleets. The absolute improvement margins are smaller, consistent with the smoother dynamics of continuous car-following that leave less room for discrete jam-breaking maneuvers.

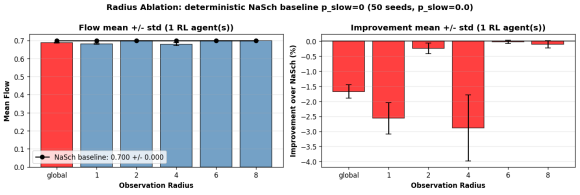
Takeaway. A local radius of $r = 5$ delivers near-optimal performance across all tested conditions while keeping the observation space fixed-size and independent of fleet size. We adopt $r = 5$ as the default for all subsequent experiments.



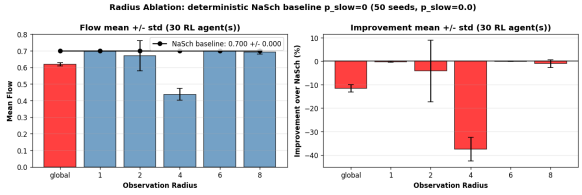
(a) $p_{\text{slow}} = 0.25$, one controlled vehicle.



(b) $p_{\text{slow}} = 0.25$, 30 controlled vehicles.



(c) $p_{\text{slow}} = 0$, one controlled vehicle.



(d) $p_{\text{slow}} = 0$, 30 controlled vehicles.

Figure 17: **Observation-radius ablation over 50 random seeds (NaSch environment).** The global setting is $r = -1$. Each panel reports the mean $\pm$ one standard deviation of flow and relative improvement over the uncontrolled baseline for radii $r \in \{-1, 1, 2, 3, 5, 7\}$. Top row: $p_{\text{slow}} = 0.25$; bottom row: $p_{\text{slow}} = 0$. Left column: one controlled vehicle; right column: 30 controlled vehicles.

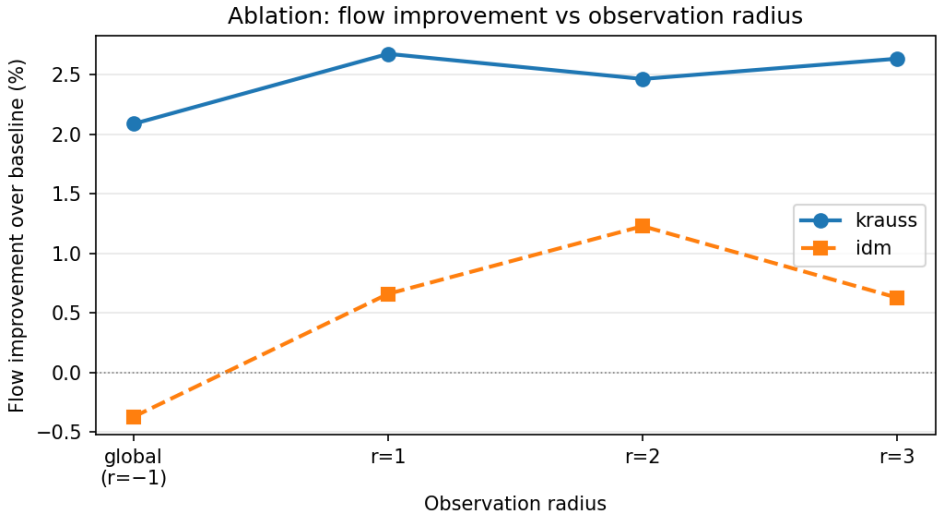


Figure 18: **Observation-radius ablation in the continuous-space environment.** Same protocol as Figure 17. The qualitative ordering is preserved; absolute margins are smaller owing to the smoother dynamics of continuous car-following models.

6.5. Feature-representation ablation

Given a fixed observation radius $r = 5$, the agent still has a choice of how to *encode* its neighborhood, a standard issue in reinforcement-learning state representation [14]. We compare two representations:

- `mean` - the neighborhood is summarized by aggregate statistics (mean gap and mean speed over the r vehicles ahead and r vehicles behind). The resulting observation vector has constant, small dimension regardless of r.
- `atomic` - each neighbor contributes an ordered pair (gap_i, v_i), preserving positional information. The observation dimension grows linearly with r.

The `mean` encoding is more compact and its statistical nature makes it robust to permutation, but it discards the relative ordering of neighbors. The `atomic` encoding retains richer positional structure at the cost of higher input dimensionality and potential overfitting. All conditions are evaluated over 50 random seeds.

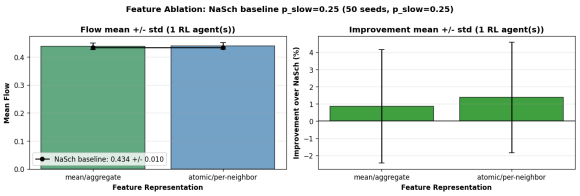
Figure 19 reports results for the same four conditions as the radius ablation.

Stochastic environment ($p_{\text{slow}} = 0.25$). Under noise, the `atomic` encoding consistently outperforms `mean` in the single-vehicle setting. The individual gap and speed of the immediately preceding vehicle carries actionable information that is smoothed away by aggregation: the agent can detect an approaching stopped car before the mean speed of the neighborhood drops. With 30 controlled vehicles the gap narrows, because agents interact frequently enough that the aggregate neighborhood statistics become more informative.

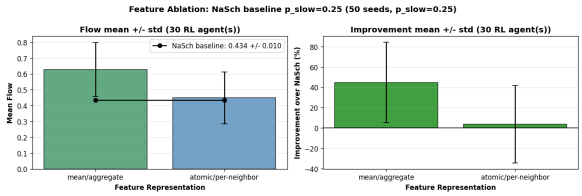
Deterministic environment ($p_{\text{slow}} = 0$). Without exogenous noise the traffic state is more homogeneous, and the two encodings perform comparably. Jams arise only from the initial condition and dissipate quickly; the agent's task reduces to maintaining a comfortable following distance, for which mean statistics are sufficient.

Variance. The standard deviation across seeds is consistently smaller for `mean`, suggesting that the lower-dimensional observation leads to a smoother optimization landscape and more reproducible policies. The `atomic` encoding achieves higher peak performance but with greater variance, indicative of a richer hypothesis class that is harder to optimize reliably.

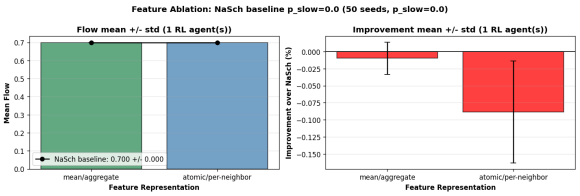
Takeaway. The `atomic` encoding is preferable when noise is present and the number of controlled agents is small - precisely the setting where individual vehicle interactions matter most. In high-penetration or low-noise regimes both encodings are roughly equivalent. Given that our primary evaluation targets the challenging single-agent stochastic setting, we adopt `atomic` as the default.



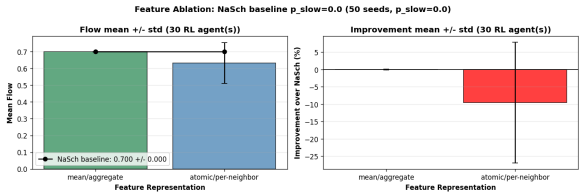
(a) $p_{\text{slow}} = 0.25$, one controlled vehicle.



(b) $p_{\text{slow}} = 0.25$, 30 controlled vehicles.



(c) $p_{\text{slow}} = 0$, one controlled vehicle.



(d) $p_{\text{slow}} = 0$, 30 controlled vehicles.

Figure 19: **Feature-representation ablation over 50 random seeds.** Observation radius fixed to $r = 5$. Each panel reports the mean $\pm$ one standard deviation of flow and relative improvement for the `mean` and `atomic` encodings. Top row: $p_{\text{slow}} = 0.25$; bottom row: $p_{\text{slow}} = 0$. Left column: one controlled vehicle; right column: 30 controlled vehicles.

7. Game Theory Approach

In this section, we provide a concise theoretical framework for modeling road traffic as a non-cooperative game. We introduce the fundamental concepts of congestion games, define equilibrium notions, and highlight the gap between individual decision-making and collective efficiency. We will focus on the understanding of the concepts rather than presenting state of the art results.

7.1. Theoretical Framework

Road traffic can be modeled as a congestion game, where multiple self-interested agents share a common infrastructure [17, 18].

A traffic game is defined by a tuple

$$\Gamma = (N, G, \{S_i\}, \{c_e\})$$

where:

- $N = \{1, \ldots, n\}$ is the set of drivers,
- $G = (V, E)$ is a directed graph (road network),
- S_i is the set of feasible routes for driver i,
- $c_e(f_e)$ is the cost (travel time) on edge e, depending on flow f_e.

The cost experienced by driver i choosing route s_i is:

$$u_i(s) = \sum_{e \in s_i} c_e(f_e)$$

where f_e is the number of drivers using edge e. Each driver minimizes their own travel time without coordinating with others, making this a non-cooperative game.
A strategy profile s^* is a Nash equilibrium if no driver can improve their travel time by unilaterally changing routes [17]:

$$u_i(s_i^*, s_{-i}^*) \leq u_i(s_i, s_{-i}^*) \quad \forall i, \forall s_i \in S_i$$

Every finite congestion game admits at least one pure Nash equilibrium through an exact potential function [17]:

$$\Phi(s) = \sum_{e \in E} \sum_{k=1}^{f_e} c_e(k)$$

Any local minimum of Φ corresponds to a Nash equilibrium.
When the number of drivers is large, traffic is modeled as a continuous flow. A flow f is a Wardrop equilibrium if all used routes have minimal cost [23]:

$$f_r > 0 \;\Rightarrow\; c_r(f) \leq c_{r'}(f) \quad \forall r'$$

If edge cost functions are continuous and non-decreasing, a Wardrop equilibrium exists. Under standard monotonicity assumptions, Wardrop equilibria can be characterized through the Beckmann transformation as a convex optimization problem [2]. The equilibrium flow solves:

$$\min_{f \geq 0} \sum_{e \in E} \int_0^{f_e} c_e(x)\, dx$$

subject to flow conservation constraints. This formulation guarantees:

- existence of a solution,
- uniqueness of edge flows when costs are strictly increasing.

The objective from a global perspective is to minimize total travel time:

$$\mathrm{SC}(f) = \sum_{e \in E} f_e \cdot c_e(f_e)$$

However, Nash and Wardrop equilibria do not generally minimize this quantity. The Price of Anarchy measures the inefficiency of selfish routing relative to the social optimum [18]:

$$\mathrm{PoA} \;=\; \frac{\max_{f^{\mathrm{NE}}} \mathrm{SC}(f^{\mathrm{NE}})}{\min_{f} \mathrm{SC}(f)} \;=\; \frac{\text{Worst Nash social cost}}{\text{Optimal social cost}} \;\geq\; 1$$

Interpretation.

- PoA $= 1 \rightarrow$ optimal system
- PoA $> 1 \rightarrow$ inefficiency due to selfish behavior

For typical congestion models (e.g. polynomial costs),the inefficiency remains bounded but can be significant.

7.2. Braess's Paradox

Braess's paradox is a fundamental phenomenon in traffic theory showing that adding capacity to a road network can increase overall congestion [3]. It arises from selfish routing behaviour and illustrates the inefficiency of Nash and Wardrop equilibria [18, 23].

Consider a network with four nodes $\{S, A, B, T\}$ and total demand N drivers traveling from S to T.

Edge costs are given by

$$t_{SA}(f) = \frac{f}{100}, \qquad t_{BT}(f) = \frac{f}{100}, \qquad t_{SB} = t_{AT} = 45, \qquad t_{AB} \approx 0.$$

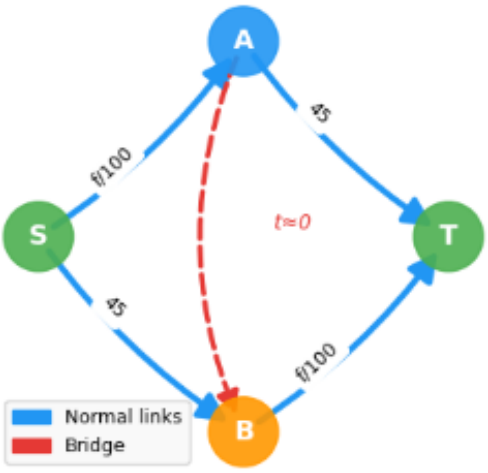


Figure 20: Graph representing a classical situation enabling the observation of Braess's Paradox.

If the link (A, B) is absent, there are two available routes:

$$P_1 : S \to A \to T, \qquad P_2 : S \to B \to T.$$

Let n_1 and n_2 denote the flows on these routes, with $n_1 + n_2 = N$. The corresponding travel times are

$$c_{P_1} = \frac{n_1}{100} + 45, \qquad c_{P_2} = 45 + \frac{n_2}{100}.$$

At Wardrop equilibrium, used routes have equal cost [23], implying

$$n_1 = n_2 = \frac{N}{2}.$$

The equilibrium travel time is therefore

$$c^*_{\text{no bridge}} = 45 + \frac{N}{200}.$$

Now suppose the link (A, B) is added, creating a third route

$$P_3 : S \to A \to B \to T.$$

Because this path avoids the constant-cost edges, each driver has an incentive to use it. At equilibrium, all drivers choose P_3, leading to flows $f_{SA} = f_{BT} = N$. The resulting travel time becomes

$$c^*_{\text{bridge}} = \frac{N}{50}.$$

The paradox occurs when the new equilibrium is worse than the original one:

$$\frac{N}{50} > 45 + \frac{N}{200}.$$

Solving this inequality yields

$$N > 3000.$$

Thus, for sufficiently large demand, adding the new road strictly increases the travel time for all users. We summarize our results in the following figure:

This phenomenon is explained by a coordination failure inherent in selfish routing [18]. Each driver minimizes their own travel time without accounting for the congestion they impose on others. As a result, the system converges to a Nash or Wardrop equilibrium that is not socially optimal. In contrast, the optimal configuration corresponds to the network without the additional link. Moreover, it is possible to observe Braess's paradox in other domains such as electronics or within complex systems like the blockchain.

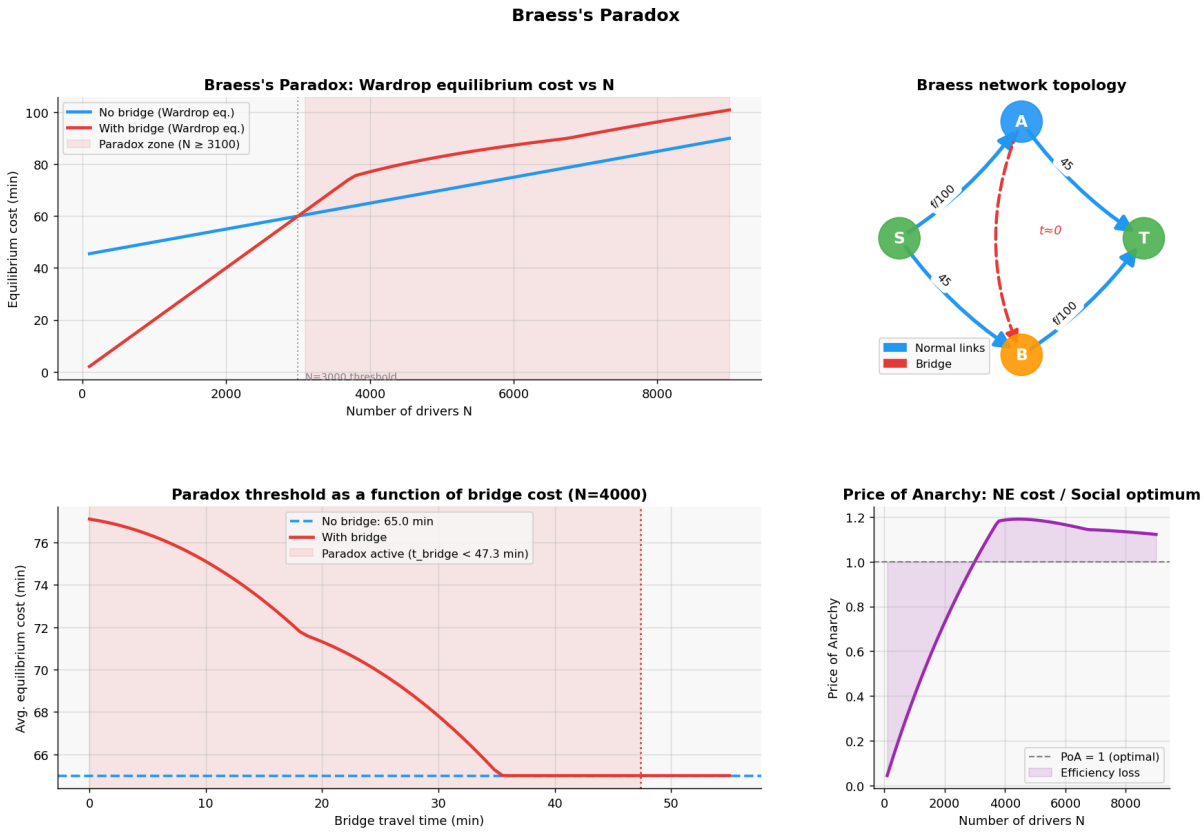


Figure 21: Braess's Paradox on our toy example.

7.3. Multi-Agent Routing Strategies

We consider a finite population of N drivers on a network $G = (V, E)$. Each driver i selects a route $s_i \in S_i$, and the resulting congestion determines the travel cost:

$$c_i(s) = \sum_{e \in s_i} t_e(f_e), \qquad f_e = |\{j : e \in s_j\}|.$$

We compare four routing strategies corresponding to different information structures, motivated by classical route-choice and selfish-routing models [18, 23]:

- **Shortest-path strategy:**
$$s_i^{\text{short}} = \arg\min_{r \in \mathcal{R}_i} \sum_{e \in r} \ell_e$$
(minimizes physical distance, ignoring congestion).

- **Fastest-path strategy:**
$$s_i^{\text{fast}} = \arg\min_{r \in \mathcal{R}_i} \sum_{e \in r} t_e(f_e)$$
(minimizes current travel time; corresponds to best-response dynamics and yields the Wardrop equilibrium as $N \to \infty$).

- **Greedy strategy:**
$$a_t = \arg\min_{v \sim u_t} t_{u_t v}(f_{u_t v})$$
(myopic decision using only local information at each step).

- **Random walk:**

$$\mathbb{P}(u \to v) \propto \frac{1}{1 + d(v, \mathrm{dst})}$$

(biased random walk toward the destination with minimal information).

A key feature of this model is the presence of a congestion externality. When a driver uses an edge e, the flow increases from f_e to $f_e + 1$, raising travel time for all users:

$$\Delta t_e = t_e(f_e + 1) - t_e(f_e).$$

This additional cost is not internalised by individual drivers, which leads to inefficient equilibria and motivates coordination mechanisms such as congestion pricing [2, 18].
We present the results obtained during our simulations. In order to work with a simple network, the road is represented as a grid.

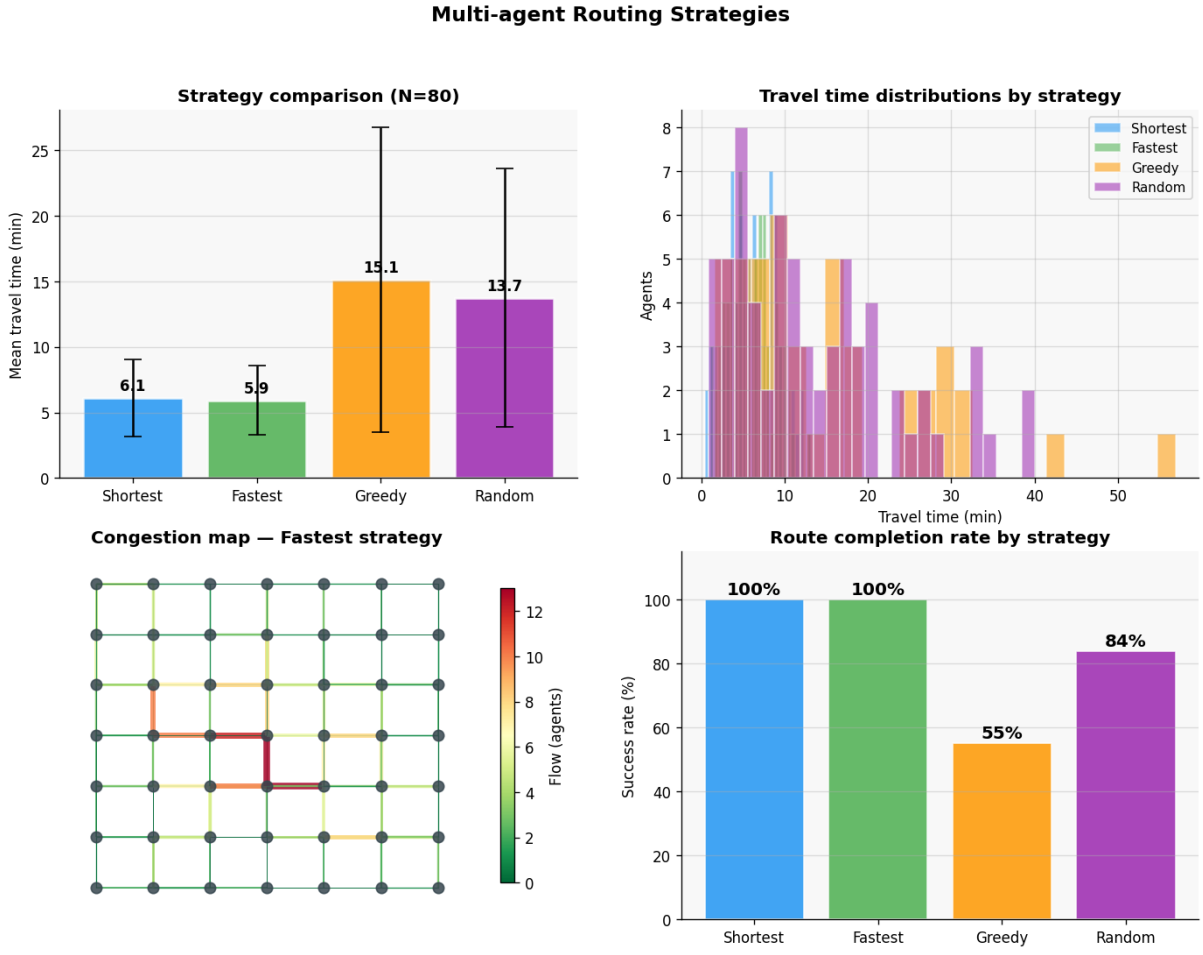


Figure 22: Comparison of the four strategies.

8. Graph-Density Optimal-Transport Model and Multi-Commodity Extension

The microscopic benchmarks established previously characterize the local car-following rules and their emergent macroscopic properties under various behavioral paradigms. To complement this bottom-up, agent-based methodology, we introduce a top-down macroscopic framework. This model describes the collective, spatiotemporal distribution of intelligent vehicles traversing a constrained urban road network. Rather than supplanting discrete-agent microscopic models,

such as the stochastic Nagel–Schreckenberg (NaSch) cellular automaton, the safe-distance Krauß model, or the continuous Intelligent Driver Model (IDM), this macroscopic layer serves a distinct theoretical purpose. It provides a physically interpretable continuum object that unifies the benchmark observables, connecting fundamental diagrams, perturbation stability, and route-choice efficiency to a strict variational principle grounded in optimal transport and Mean Field Game (MFG) theory [1, 12]. We present both a baseline single-commodity formulation to establish the underlying thermodynamics and an advanced multi-commodity MFG extension that captures the strategic interactions of heterogeneous origin–destination groups.

8.1. Geometry, Network Topology, and State Variables

The urban road network is formalized rigorously as a metric directed graph $G = (V, E)$. In this topology, edges $e \in E$ represent unidirectional road segments characterized by a physical length ℓ_e, while vertices $v \in V$ correspond to intersections governing mass transfer. This precise geometric embedding is critical: the inherent graph distance metric $d_{\text{graph}}(x, y)$ respects the complex topology of one-way streets, mandatory turn restrictions, and network discontinuities, distinguishing it fundamentally from standard Euclidean spaces.

The spatial domain is configured into two distinct experimental setups to isolate specific routing phenomena. The baseline single-commodity grid consists of a Manhattan-style network comprising 21×21 blocks. This yields exactly 484 intersections and 924 directed road segments. Within this geometry, fifteen boundary nodes operate as fluid entrances, while a singular central node functions as the absorbing target. The multi-commodity grid, designed to analyze inter-group conflict, utilizes a 10×10 Manhattan layout. It features 121 intersections indexed by spatial coordinates (x, y) and 440 directed edges. The reduction in spatial resolution from the baseline model is deliberately chosen to maintain computational tractability while preserving the topological complexity necessary for analyzing multi-group interactions.

In the multi-commodity formulation, nine public squares are designated at coordinates forming a regular 3×3 lattice within the grid. These nodes function mathematically as absorbing sinks where probability mass permanently exits the system upon arrival. Physically, these represent high-capacity parking facilities or concentrated drop-off zones within a central business district. The remaining 112 intersections are partitioned based on their functional capacity. Seventy-two interior nodes are classified as signalized intersections, where traffic control mechanisms restrict the effective flow capacity to $\rho_{\max}^{\text{sig}} = 0.40\rho_{\max}$. This reduction factor explicitly models the fraction of available green time and the requisite lost time during phase transitions. Conversely, the forty boundary intersections forming the network perimeter are assumed to operate without signal interruption, maintaining a full geometric capacity of $\rho_{\max} = 1.0$ vehicle per unit cell, corresponding approximately to 133 vehicles per kilometer per lane at standard jam densities.

The instantaneous traffic state is defined by a nonnegative, continuous density function defined across the edges and vertices of the graph. To capture origin-destination heterogeneity, the multi-commodity model tracks $k = 1, \ldots, 8$ distinct vehicle groups, each governed by an independent density field $\rho_k(x, t)$. This commodity disaggregation provides critical visibility into group-specific routing behaviors, dynamic travel times, and spatial conflict zones that remain fundamentally unobservable in aggregate density models. The total instantaneous network density is defined simply as the linear superposition $\rho_{\text{total}}(x, t) = \sum_{k=1}^{8} \rho_k(x, t)$. The system's state space is constrained by an admissible set $\mathcal{K}$ that strictly enforces the macroscopic safety-gap requirement:

$$\mathcal{K} = \{\rho \geq 0 : \rho_{\text{total}}(x) \leq \rho_{\max}(x) \text{ for almost every } x \in G\}.$$

8.2. Thermodynamic Free Energy Functional and Game-Theoretic Foundations

The temporal evolution of the traffic network is governed completely by the gradient flow of a global free energy functional. This mathematical architecture unifies local entropy, external

potential fields, and nonlocal density interactions within a cohesive thermodynamic formalism [1]. For the multi-commodity system, the global functional decomposes into group-specific energies and a shared structural capacity penalty:

$$\mathcal{F}[\rho_1, \dots, \rho_8] = \sum_{k=1}^{8} \mathcal{F}_k[\rho_k | \rho_{\text{total}}] + \mathcal{F}_{\text{capacity}}[\rho_{\text{total}}]$$

Each individual commodity's free energy contribution is defined by three physically distinct integral terms over the one-dimensional Hausdorff measure dH^1 of the network:

$$\mathcal{F}_k = \epsilon \int_G \rho_k \log(\rho_k + \eta)\, dH^1 + \int_G \Phi_k(x)\rho_k(x)\, dH^1 + \frac{\chi}{2} \int_G \rho_k(x)[K * \rho_{\text{total}}](x)\, dH^1$$

Furthermore, the network-level capacity limit is enforced via a differentiable penalty function, which serves as a computationally tractable approximation to the rigid indicator functional $I_{\mathcal{K}}(\rho)$:

$$\mathcal{F}_{\text{capacity}} = \frac{\beta}{2} \int_G [\rho_{\text{total}}(x) - \rho_{\max}(x)]_+^2\, dH^1$$

The physical interpretation of these integral components defines the macroscopic behavioral physics of the model. The entropy term, scaled by $\epsilon = 0.02$, originates directly from the Boltzmann formulation. Its first variation produces Fickian diffusion with mobility proportional to $\epsilon\rho_k$. Physically, this mechanism suppresses the formation of unphysical discontinuous shock-waves typical of purely hyperbolic Lighthill–Whitham–Richards (LWR) conservation laws [13]. Behaviorally, it models the irreducible stochasticity in driver routing choices and acceleration profiles. Even within a homogeneous cohort sharing identical origins and destinations, drivers will disperse across slightly varying trajectories and speed distributions. The minute regularization parameter $\eta = 10^{-9}$ is required purely to prevent logarithmic singularities in regions of absolute vacuum.

The destination potential Φ_k acts as a static, external force field drawing group k toward its specific target sink s_k. This field is constructed using the normalized squared graph distance:

$$\Phi_k(x) = \frac{1}{2} \frac{d_{\text{graph}}(x, s_k)^2}{\max_{y \in V} d_{\text{graph}}(y, s_k)}$$

The reliance on a squared-distance formulation is intrinsically linked to the geometry of optimal transport with a quadratic Wasserstein cost function. It ensures that the marginal disutility of deviating from the optimal shortest path scales linearly with the magnitude of the deviation, heavily penalizing extreme detours while permitting minor spatial variances.

The congestion interaction parameter, $\chi = 0.45$, introduces a nonlocal repulsive force coupling all vehicle classes. The convolution integral evaluates the interaction kernel against the total network density, penalizing configurations that promote excessive vehicular clustering. This induces an anticipatory route diversification mechanism, compelling simulated drivers to select alternative trajectories prior to the exhaustion of hard spatial capacity. This mirrors Wardrop's first principle of equilibrium applied in a dynamic continuum context [23]. The kernel itself aggregates a regional Gaussian spread (250 meter range), a bidirectional suppression penalty to model the degradation of intersection throughput under heavy counter-flowing traffic, and an localized penalty isolating the specific queueing externalities native to signalized intersections.

Finally, the capacity penalty coefficient $\beta = 30.0$ establishes a soft operational barrier. As $\beta \to \infty$, the strict admissibility constraint $\mathcal{K}$ is recovered. However, maintaining a finite, strictly positive β permits slight volumetric compressibility. This mathematical relaxation closely approximates real-world traffic flows, wherein vehicle fleets possess a distribution of physical lengths and operators exhibit variance in acceptable safety gaps under extreme congestion pressure.

8.3. Evolution Equations, the JKO Scheme, and Mean Field Games

The local chemical potential governing the motion of group k is derived through functional differentiation of the total free energy:

$$\mu_k(x,t) = \frac{\delta \mathcal{F}}{\delta \rho_k}(x,t) = \epsilon \log(\rho_k + \eta) + \epsilon \frac{\rho_k}{\rho_k + \eta} + \Phi_k(x) + \chi[K * \rho_{\text{total}}](x) + \beta[\rho_{\text{total}} - \rho_{\max}]_+$$

Assuming the dilute limit where $\rho_k \gg \eta$, the entropic contribution simplifies, recovering the standard diffusive gradient $\epsilon \nabla \log \rho_k$. The corresponding macroscopic velocity field for each commodity is defined as the negative graph-gradient of this chemical potential. To prevent arbitrary velocity magnitudes driven by steep potential gradients near saturated bottlenecks, the velocity is projected onto the tangent space of the directed edges and bounded by an urban free-flow speed constraint $v_{\max} = 2.0$ m/s via a hyperbolic tangent saturation function. The transport of the probability mass is then dictated strictly by the continuity equation $\partial_t \rho_k + \nabla_G \cdot (\rho_k v_k) = 0$, ensuring exact conservation of vehicles at all network vertices via Kirchhoff flux matching.

This coupled dynamic system exhibits a profound structural duality. As demonstrated by Jordan, Kinderlehrer, and Otto (JKO), the temporal evolution constitutes a mathematically exact gradient flow of $\mathcal{F}$ within the Wasserstein-2 metric space of probability measures [8]. Under the JKO time-discretization scheme utilizing a time step τ, the density sequence updates via:

$$\rho^{n+1} = \arg\min_{\rho \in \mathcal{K}} \left\{ \frac{1}{2\tau} W_{2,G}^2(\rho, \rho^n) + \mathcal{F}(\rho) \right\}$$

This reveals that the traffic mass perpetually reorganizes to optimally balance the reduction of global thermodynamic free energy against the geometric resistance to mass transport defined by the Wasserstein metric.

Simultaneously, this gradient flow represents the limiting continuous solution to a specific Mean Field Game [12]. In this interpretation, each microscopic agent minimizes a stochastic running cost composed of kinetic energy, distance potential, and nonlocal congestion penalties. In the infinite-agent mean-field limit, the equilibrium is described by the coupled Hamilton–Jacobi–Bellman (HJB) and Fokker–Planck (FP) differential equations. The congestion pressure mathematically manifests as the Lagrange multiplier enforcing the density constraint across the network, dynamically pricing the spatial scarcity of the road infrastructure.

8.4. Numerical Discretization and Network Evaluation

To preserve both exact mass conservation and strict capacity bounds, the simulation employs a conservative, two-scale operator-splitting algorithm. At the micro-scale ($\Delta t = 0.5$ s), density advection is resolved using a robust upwind donor-cell finite volume scheme. Fluxes between adjacent discrete 100-meter cells are rigorously clipped to satisfy both the Courant–Friedrichs–Lewy (CFL) stability criterion and the static capacity bounds. At intersections, an intermediate vertex buffer collects all incoming discrete fluxes and distributes them proportionally according to downstream edge capacities, precluding any artificial numerical creation or destruction of mass.

At the macro-scale ($\tau = 5.0$ s), the density fields undergo JKO relaxation toward the local free-energy minimum. Simultaneously, vehicles occupying designated sink nodes are systematically removed using a continuous exponential decay function $\rho_k(s, t+\tau) = \rho_k(s,t) e^{-\gamma_{\text{sink}} \tau}$. Setting the service rate to $\gamma_{\text{sink}} = 0.05$ s^{-1} elegantly captures the finite absorption bandwidth of terminal parking facilities, inducing realistic upstream queuing queues characteristic of heavily saturated destination zones.

8.5. Analysis of Dynamics and Emergent Network Phenomena

The macroscopic model bridges the gap between local driving rules and network-wide phase transitions. Figure **??** contrasts the fundamental diagrams generated by discrete-agent benchmarks

(NaSch, Krauß, IDM) with the perturbation response of the continuous model. The microscopic rules yield distinct capacity regimes, from the sharp stochastic collapse of the NaSch model to the smooth deterministic flow of the IDM. The macroscopic free energy parameters β and χ are explicitly calibrated to align the continuum model's equilibrium manifold with the concave envelope of these micro-relations. Furthermore, the macroscopic continuum successfully models finite-time recovery from localized braking perturbations, a phenomenon typically reserved for high-fidelity car-following algorithms, demonstrating that the optimal-transport framework captures the requisite fluid-dynamic compressibility.

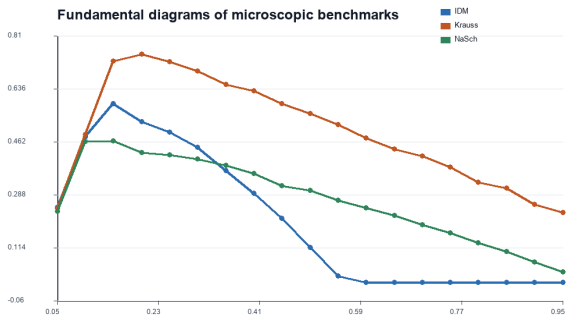


Figure 23: Fundamental diagrams derived from discrete-agent microscopic models (NaSch, Krauß, IDM), illustrating variations in critical density and capacity drop due to differing stochastic and deterministic update rules.

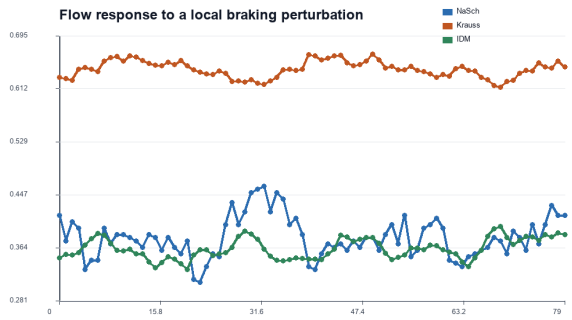


Figure 24: Time-series flow response following a localized braking perturbation. The continuum macroscopic formulation successfully reproduces the damped oscillatory recovery characteristic of bounded driver reaction times.

The strategic routing behavior intrinsic to the MFG formulation naturally circumvents classical routing pathologies. As evidenced in Figure **??**, the optimal transport formulation avoids the Braess Paradox [3]. Because the underlying dynamics seek to minimize the global free energy rather than allowing localized selfish optimization, the addition of a seemingly optimal central link does not degrade total system throughput. The entropy-driven dispersion actively prevents the over-concentration of mass that precipitates the paradox under standard Wardrop equilibrium conditions. In parallel, the macroscopic representation provides a powerful analytical lens for mixed-autonomy scenarios. We observe a highly nonlinear, non-monotonic flow response with respect to the market penetration of autonomous agents. The residual-control diagnostic confirms that conservative autonomous spacing algorithms deployed at low penetration rates actively degrade the cohesive platoon structures formed by human drivers, dictating that net system benefits are exclusively realized when autonomous fleets exceed a forty percent saturation threshold.

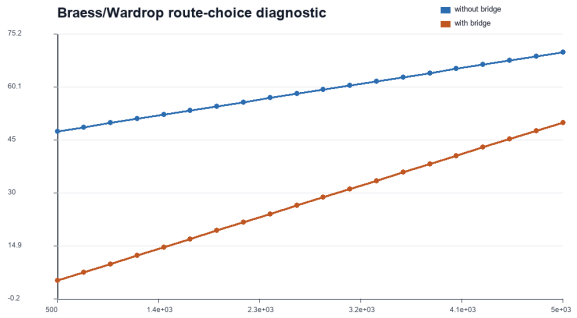


Figure 25: Evaluation of the Braess diagnostic network. The free energy gradient flow fundamentally resists the paradox by dynamically penalizing the over-concentration of density on theoretical shortcuts.

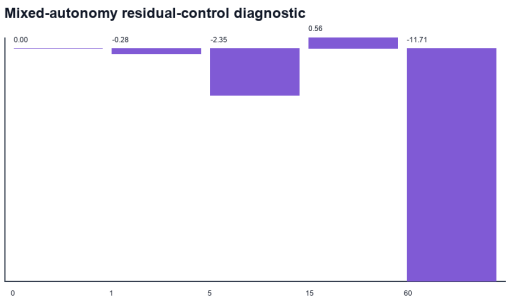


Figure 26: Mixed-autonomy residual-control flow gain. The non-monotonic trajectory illustrates the interference pattern between stochastic human drivers and deterministic autonomous agents at low penetration thresholds.

The deployment of the multi-commodity setup exposes critical higher-order spatiotemporal patterns. Figure **??** presents the Macroscopic Fundamental Diagram (MFD) extracted across the entire operational network alongside an analysis of bidirectional phase-locking. The network MFD successfully generates an emergent hysteresis loop, the operational capacity drop. During the dissipation phase, flow rates corresponding to specific network densities are demonstrably lower than identical densities observed during the initial loading phase. This discrepancy arises physically because residual high-density gridlock pockets persist in the network core long after the peripheral arterials have evacuated, severely depressing global flow efficiency.

Simultaneously, the model resolves complex localized intersection dynamics without explicit microscopic modeling. The bidirectional suppression kernel produces spontaneous phase-locking on highly contested corridors. Competing vehicle groups traveling in opposite directions rarely achieve high simultaneous densities within the identical block. Instead, the nonlinear repulsive forces organize the traffic into alternating, temporally offset pulses. This phenomenon highlights the capability of the macroscopic continuum to organically resolve mutual spatial obstruction into coordinated, oscillating flow regimes, mirroring self-organized signal phasing.

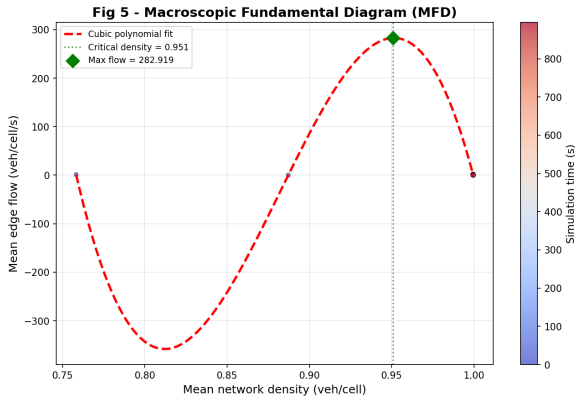


Figure 27: Network-wide Macroscopic Fundamental Diagram (MFD). The explicit separation of the loading and unloading curves clearly establishes the emergent spatial hysteresis and persistent capacity degradation.

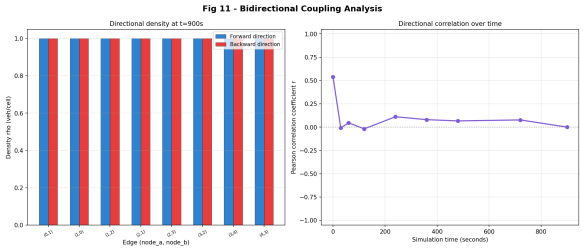


Figure 28: Bidirectional phase-locking resulting from the nonlocal density interaction kernel. Opposing macroscopic flows spontaneously self-organize into alternating temporal pulses to maximize local throughput.

While the free-energy gradient framework successfully bounds the density strictly within theoretical capacity constraints and accurately tracks exact mass conservation to machine precision, its limitations must be clearly defined. The uniform spatial homogenization over 100-meter blocks mathematically eliminates the resolution of distinct lane-changing dynamics and precise vehicle-pedestrian interactions. Furthermore, the reliance on a static signalization parameter ($\rho_{\max}^{\text{sig}}$) prevents the analysis of reactive, real-time adaptive signal control policies, suggesting the model represents a conservative lower bound on the maximum theoretical efficiency of the evaluated network geometry.

9. Discussion

The experiments support a multi-scale interpretation of congestion. Traffic breakdown is not a single mechanism but a family of interaction driven regimes that can arise through stochastic braking, safety constrained car following, deterministic instability, selfish routing on a shared network, or global thermodynamic balancing.

At the microscopic level, the three benchmarked models generate similar macroscopic regimes through different causal mechanisms. In the Nagel-Schreckenberg (NaSch) model, random decelerations create local perturbations that grow into backward propagating phantom jams near the critical density, producing fragmented spacetime patterns with sharp boundaries [15]. In the Krauss model, stochasticity is filtered through a safe speed constraint, yielding smoother trajectories but a strong capacity penalty. Conservative safety margins reduce collisions at the cost of throughput [10]. In the IDM, congestion can emerge from deterministic nonlinear feedback alone, so desired time headway, acceleration, and minimum spacing act as stability controls rather than purely behavioral parameters [20].

The observer suite confirms that aggregate flow alone is insufficient to characterize traffic quality. A configuration can achieve high mean velocity while producing high velocity variance, frequent emergency braking, or dangerously low time to collision values. Conversely, conservative driving reduces safety critical events but lowers throughput. The phase diagrams further show that congestion onset depends jointly on density and behavioral parameters. There is no universal critical density independent of stochasticity, reaction time, desired headway, or maximum velocity.

The reinforcement learning experiments connect this descriptive analysis to a control problem. Learned policies improve traffic flow by locally modifying vehicle behavior, with performance growing with penetration rate. Spacetime diagrams suggest that controlled vehicles act as moving dampers that absorb perturbations and disrupt stop and go waves, consistent with the mixed autonomy literature [11, 22, 24]. The ablation results show that global observations provide the largest advantage when controlled vehicles are sparse, but moderate local sensing radii become competitive once enough controlled agents are distributed across the ring. This suggests that deployable traffic control does not necessarily require full state observation, provided the controlled fleet has sufficient spatial coverage. Cross model transfer results further indicate that learned policies retain part of their effectiveness across models, though with degradation that reflects differences in the underlying dynamics.

The game theoretic component highlights a distinct source of congestion found in inefficient route choice. While the microscopic benchmark isolates car following interactions on a ring, congestion games and Wardrop equilibria describe how selfish drivers distribute themselves across a network [2, 17, 23]. Simulating four routing strategies on a grid network, we find that as demand grows, the gap between individual and socially optimal routing widens. Beyond a specific demand threshold, Braess's paradox emerges [3, 18]. Drivers collectively overload the new shortcut and everyone's travel time strictly increases. Real urban congestion combines both mechanisms studied in this paper. Vehicles interact locally through acceleration and braking, but also strategically through route choices that determine where density accumulates.

The introduction of the macroscopic optimal transport framework serves as a unifying theoretical bridge. While the microscopic models detail the mechanics of congestion and the game theoretic models highlight routing failures, the continuum model explains the organizing principles driving these states. By minimizing a global free energy functional, the model organically reproduces the concave fundamental diagrams observed in the microscopic simulations. It also reveals emergent spatiotemporal phenomena, such as network wide capacity hysteresis and bidirectional phase locking. Because the density evolution follows a Wasserstein gradient flow [1, 8], it naturally resists the over concentration that causes Braess's paradox, confirming that systemic optimal transport can decentralize the coordination failures inherent to selfish routing.

Several limitations should be acknowledged. The microscopic benchmark uses a periodic single lane ring, which isolates spontaneous jam formation but excludes lane changing, intersections, traffic lights, and open boundary demand fluctuations. The reinforcement learning controller is evaluated in simulation and remains sensitive to reward design, observation radius, and fleet penetration. The game theoretic experiments use stylized networks rather than calibrated urban demand. Additionally, the macroscopic continuum relies on a coarse spatial discretization and static capacity parameters, homogenizing fine grained lane changes and adaptive signal controls. These limitations do not invalidate the framework, but they delimit the claims. This paper provides a controlled comparative study of traffic interactions, not a calibrated operational simulator.

Overall, the four layers of this study form a coherent picture of how congestion forms and persists. Microscopic rules determine how local perturbations grow into jams, learned rl controllers reshape these interactions through targeted interventions, selfish routing explains why inefficient network states are stable even when infrastructure improves, and the optimal transport continuum provides the underlying thermodynamic principles that govern the entire system.

10. Conclusion

We presented a multi scale study of traffic congestion as an interaction driven phenomenon, organized around four complementary modeling approaches.

At the microscopic level, we compared the Nagel and Schreckenberg, Krauss, and Intelligent Driver models under a common ring road benchmark using shared density sweeps, parameter sweeps, and observer metrics. All three reproduce similar qualitative regimes including free flow, stop and go waves, and congestion through distinct mechanisms. These include stochastic braking and exclusion in NaSch [15], safety constrained stochastic dynamics in Krauss [10], and nonlinear deterministic car following instability in IDM [20]. A central lesson is that throughput alone is insufficient to evaluate traffic performance. Velocity variance, local density heterogeneity, jam morphology, emergency braking, gap zero events, and time to collision all reveal trade offs between efficiency, stability, and safety that the fundamental diagram obscures.

We then studied reinforcement learning control in a mixed autonomy setting, introducing controlled vehicles at increasing penetration rates from a single agent up to a significant fraction of the fleet. Learned policies can improve aggregate flow by modifying local vehicle interactions, with performance growing with penetration rate. Observation ablations show that global information matters most when controlled vehicles are sparse, while local sensing becomes sufficient once they are distributed throughout the system. Cross model transfer experiments further show that policies trained on one model can retain part of their effectiveness on the others without retraining, though with some degradation that depends on how different the underlying dynamics are.

We also approached congestion from a game theoretic angle. Formalizing road traffic as a congestion game and studying Nash and Wardrop equilibria [17, 23], the Beckmann potential [2], and the Price of Anarchy [18], we showed that selfish routing alone is enough to produce inefficient network states. Simulating four routing strategies on a grid network, we identified a

demand threshold beyond which Braess's paradox [3] emerges. Adding a road strictly worsens everyone's travel time because drivers collectively overload the new shortcut. This confirms that congestion is not only a car following instability but also a coordination failure that persists even when infrastructure improves.

To integrate these discrete interactions into a continuous mathematical framework, we introduced a macroscopic optimal transport model. Grounded in Mean Field Game theory [12] and Wasserstein gradient flows [1], this multi commodity formulation tracks heterogeneous vehicle groups minimizing a global free energy functional. By balancing entropic diffusion, destination potentials, and nonlocal congestion penalties, the continuum model successfully recovers empirical fundamental diagrams and emergent spatial behaviors like network wide capacity hysteresis. Furthermore, it mathematically circumvents the coordination failures of selfish routing, demonstrating how systemic optimal transport naturally mitigates congestion.

Taken together, these four layers form a coherent picture. Microscopic models explain how local braking rules generate jams. Reinforcement learning tests whether small interventions can dissolve them. Game theory explains why selfish routing sustains congestion at the network level even in the absence of any dynamical instability. Finally, the optimal transport continuum establishes the overarching thermodynamic principles that govern the entire system. Future work could extend the benchmark to multilane roads, lane changing rules, and open boundary inflows, calibrate the models on real urban networks, and explore integrated control mechanisms that combine reinforcement learning with congestion pricing or cooperative routing.

10.1. Future Work

Several natural extensions follow from this work. We began exploring multilane dynamics with explicit lane-changing rules but could not obtain conclusive results within the project timeline, making this the most immediate next step. Open-boundary networks with stochastic inflows would also allow us to study bottleneck activation under more realistic demand conditions.

For the reinforcement learning component, the main open questions concern deployability under sensor noise and actuation latency, and whether multi-agent coordination through vehicle-to-vehicle communication could further improve wave dissipation. On the game-theoretic side, extending to dynamic routing with congestion pricing would provide a concrete mechanism to prevent coordination failures like Braess's paradox.

The longer-term goal would be a coupled simulator where the macroscopic layer governs network-wide route choice while the microscopic models resolve local vehicle interactions.

References

[1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Birkhäuser, 2008.

[2] Martin Beckmann, C. B. McGuire, and Christopher B. Winsten. *Studies in the Economics of Transportation*. Yale University Press, 1956.

[3] Dietrich Braess. Über ein paradoxon aus der verkehrsplanung. *Unternehmensforschung*, 12:258–268, 1968.

[4] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, pages 226–231, 1996.

[5] Bruce D. Greenshields. A study of traffic capacity. *Proceedings of the Highway Research Board*, 14:448–477, 1935.

[6] John C. Hayward. Near-miss determination through use of a scale of danger. Technical Report TTSC 7115, Pennsylvania State University, 1972.

[7] Dirk Helbing. Traffic and related self-driven many-particle systems. *Reviews of Modern Physics*, 73(4):1067–1141, 2001.

[8] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the fokker–planck equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998.

[9] Boris S. Kerner and Hubert Rehborn. Experimental properties of phase transitions in traffic flow. *Physical Review Letters*, 79(20):4030–4033, 1997.

[10] Stefan Krauß. *Microscopic Modeling of Traffic Flow: Investigation of Collision Free Vehicle Dynamics*. PhD thesis, Universität zu Köln, 1998.

[11] Abdul Rahman Kreidieh, Cathy Wu, and Alexandre M. Bayen. Dissipating stop-and-go waves in closed and open networks via deep reinforcement learning. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 1475–1480, 2018.

[12] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese Journal of Mathematics*, 2:229–260, 2007.

[13] Michael J. Lighthill and Gerald B. Whitham. On kinematic waves. ii. a theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A*, 229(1178):317–345, 1955.

[14] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, 2015.

[15] Kai Nagel and Michael Schreckenberg. A cellular automaton model for freeway traffic. *Journal de Physique I*, 2(12):2221–2229, 1992.

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

[17] Robert W. Rosenthal. A class of games possessing pure-strategy nash equilibria. *International Journal of Game Theory*, 2(1):65–67, 1973.

[18] Tim Roughgarden and Éva Tardos. How bad is selfish routing? *Journal of the ACM*, 49(2):236–259, 2002.

[19] Yuki Sugiyama, Minoru Fukui, Macoto Kikuchi, Katsuya Hasebe, Akihiro Nakayama, Katsuhiro Nishinari, Shin-ichi Tadaki, and Satoshi Yukawa. Traffic jams without bottlenecks—experimental evidence for the physical mechanism of the formation of a jam. *New Journal of Physics*, 10(3):033001, 2008.

[20] Martin Treiber, Ansgar Hennecke, and Dirk Helbing. Congested traffic states in empirical observations and microscopic simulations. *Physical Review E*, 62(2):1805–1824, 2000.

[21] Martin Treiber, Ansgar Hennecke, and Dirk Helbing. Congested traffic states in empirical observations and microscopic simulations. *Physical review E*, 62(2):1805, 2000.

[22] Eugene Vinitsky, Aboudy Kreidieh, Luc Le Flem, Nishant Kheterpal, Kathy Jang, Cathy Wu, Fangyu Wu, Richard Liaw, Eric Liang, and Alexandre M. Bayen. Benchmarks for reinforcement learning in mixed-autonomy traffic. In *Proceedings of the 2nd Conference on Robot Learning*, volume 87 of *Proceedings of Machine Learning Research*, pages 399–409, 2018.

[23] John Glen Wardrop. Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*, 1(3):325–362, 1952.

[24] Cathy Wu, Abdul Rahman Kreidieh, Kanaad Parvate, Eugene Vinitsky, and Alexandre M. Bayen. Flow: A modular learning framework for mixed autonomy traffic. *IEEE Transactions on Robotics*, 38(2):1270–1286, 2022.

A. Theoretical Derivations

A.1. Critical Density in the Deterministic NaSch Model

In the Nagel-Schreckenberg model, the physical space is discrete. Each vehicle occupies exactly one cell of length $l = 1$. In the deterministic limit ($p = 0$), the maximum flow is achieved when every vehicle is moving at the maximum velocity $v_{\max}$ with the minimum possible safe gap.

Because vehicles move $v_{\max}$ cells per time step, the spatial gap between the rear of the leading vehicle and the front of the following vehicle must be exactly $v_{\max}$ cells to prevent any necessary braking.

Therefore, the total spatial footprint of one vehicle and its required forward gap is $v_{\max} + 1$ cells. The critical density ρ_c, representing the maximum density before free flow breaks down, is the inverse of this footprint:

$$\rho_c = \frac{1}{v_{\max} + 1}$$

For densities $\rho > \rho_c$, the available space per vehicle is strictly less than $v_{\max} + 1$. This means at least one vehicle must travel at a speed $v < v_{\max}$, which triggers a backward-propagating wave of congestion.

A.2. Equilibrium Fundamental Diagram of the IDM

The Intelligent Driver Model defines the acceleration $\dot{v}$ of a vehicle as:

$$\dot{v} = a \left[1 - \left(\frac{v}{v_0} \right)^{\delta} - \left(\frac{s^*(v, \Delta v)}{s} \right)^2 \right]$$

where the desired dynamic gap is:

$$s^*(v, \Delta v) = s_0 + vT + \frac{v \Delta v}{2\sqrt{ab}}$$

To find the theoretical steady-state fundamental diagram, we assume homogeneous, stationary traffic. In this state, all vehicles travel at the same constant equilibrium velocity v. Consequently, acceleration is zero ($\dot{v} = 0$) and the velocity difference between consecutive vehicles is zero ($\Delta v = 0$).

Substituting these steady-state conditions into the IDM equations simplifies the desired gap to:

$$s^*(v, 0) = s_0 + vT$$

Setting the acceleration equation to zero yields:

$$0 = a \left[1 - \left(\frac{v}{v_0} \right)^{\delta} - \left(\frac{s_0 + vT}{s_e(v)} \right)^2 \right]$$

Solving for the equilibrium gap $s_e(v)$ gives:

$$s_e(v) = \frac{s_0 + vT}{\sqrt{1 - \left(\frac{v}{v_0} \right)^{\delta}}}$$

The macroscopic density ρ is the inverse of the gross distance between vehicles. Assuming a vehicle length l, the equilibrium density as a function of velocity is:

$$\rho_e(v) = \frac{1}{s_e(v) + l}$$

This parametric relationship analytically defines the equilibrium flow-density relation for the IDM, where flow is defined as $J_e(v) = \rho_e(v) v$.

B. Additional Figures

C. Detailed Macroscopic and Safety Metrics: Nasch model

This section provides the detailed multi-panel metric tracking for the Nagel–Schreckenberg benchmark, expanding upon the macroscopic summary provided in the main text.

Figure 29 illustrates the impact of the stochastic randomization parameter (p_{slow}) across six core dimensions. The sharp inflection points in density dispersion, emergency braking, and gap-zero events strongly correlate with the phase transition boundary observed in the macroscopic phase diagram.

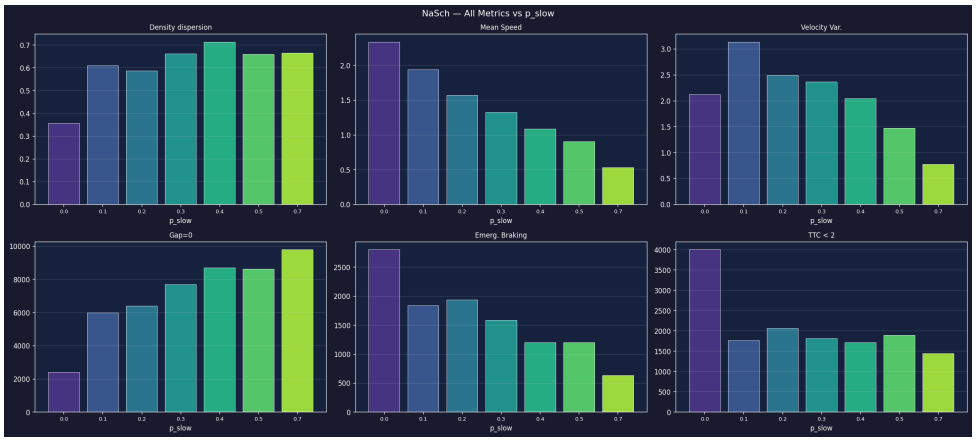


Figure 29: **NaSch Core Metrics (p_{slow})**: Sensitivity of mean velocity, velocity standard deviation, local density dispersion, gap-zero events, emergency braking count, and TTC to increasing stochastic noise. The sharp increase in density dispersion and safety events marks the onset of congested flow.

Figure 30 isolates the effect of the physical speed limit ($v_{\max}$) on the system's safety profile. The data demonstrates that the severity of traffic waves scales non-linearly with maximum velocity, resulting in degraded Time-to-Collision (TTC) and heightened emergency braking rates when fast-moving vehicles encounter standing congestion.

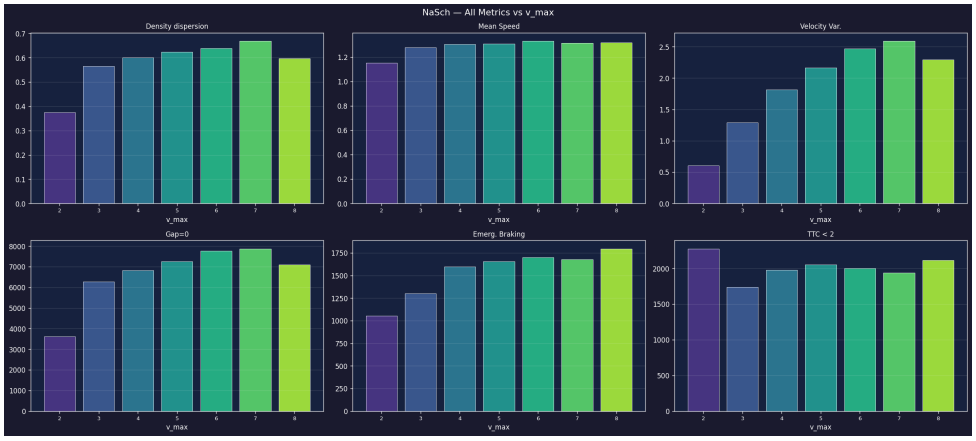


Figure 30: **NaSch Core Metrics ($v_{\max}$)**: Sensitivity of core traffic metrics to the speed limit. While higher $v_{\max}$ increases free-flow capacity, it also amplifies safety-critical events at jam boundaries.

D. Detailed Macroscopic and Safety Metrics: Krauss Model

This section provides the detailed multi-panel metric tracking for the Krauss benchmark, complementing the macroscopic phase analysis in the main text.

Figure 31 details the system's sensitivity to the noise amplitude (ϵ). While the mean velocity is robust to moderate noise, safety metrics (gap-zero events, emergency braking, and TTC) degrade progressively as ϵ grows, reflecting the continuous accumulation of micro-perturbations.

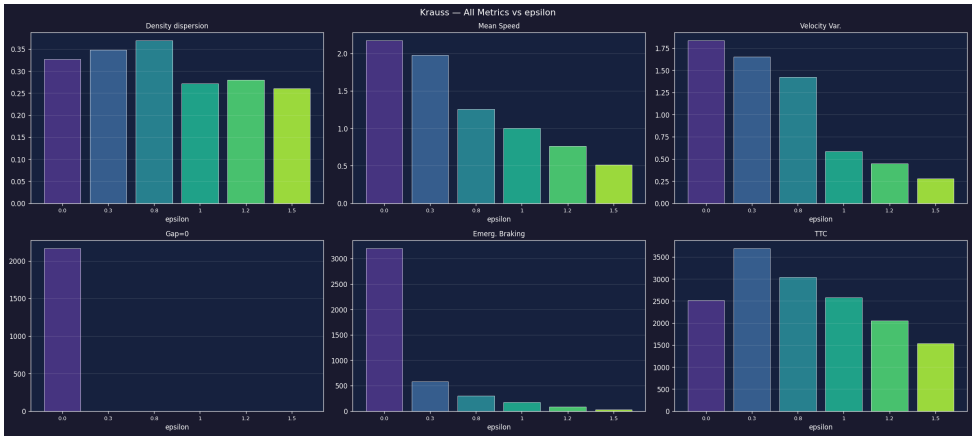


Figure 31: **Krauss Core Metrics (ϵ)**: Sensitivity of traffic metrics to noise amplitude. Mean velocity is robust to moderate noise, while safety metrics degrade progressively as ϵ increases.

Figure 32 highlights the severe capacity penalty incurred by high reaction times (τ). While high τ effectively eliminates safety-critical events by enforcing conservative headways, it substantially suppresses mean velocity and throughput across all density regimes.

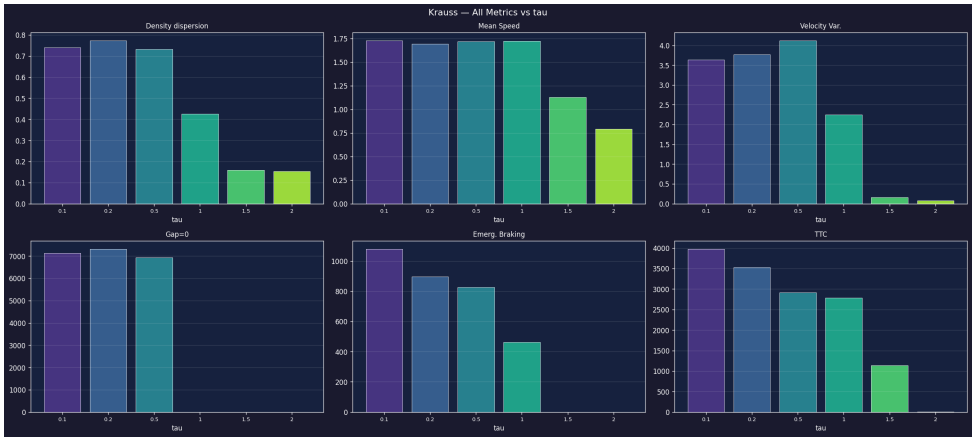


Figure 32: **Krauss Core Metrics (τ)**: Sensitivity of traffic metrics to reaction time. While high τ improves safety metrics by enforcing large headways, it substantially reduces road capacity and mean velocity.

E. Detailed Macroscopic and Safety Metrics: IDM

This section provides the detailed multi-panel metric tracking and jam morphological analysis for the Intelligent Driver Model (IDM) benchmark.

Impact of Time Headway (T) and Noise (ϵ)

Figure 33 highlights the severe impact of deterministic instability at low T, where emergency braking peaks. Figure 34 confirms that at critically low T, the system collapses into a single large jam. Figure 35 and Figure 36 contrast this with noise-induced instabilities, showing that stochastic jams are significantly more spatially fragmented. Flux-based detection proves vastly superior to density-based mapping in resolving these individual clusters.

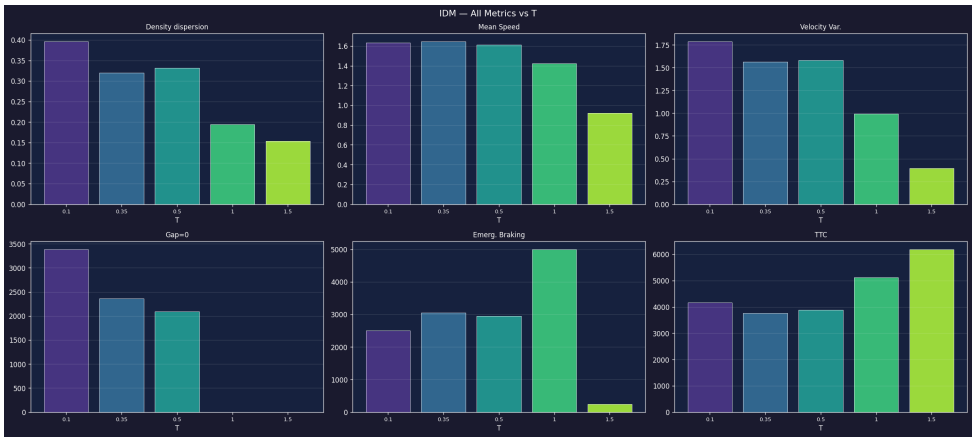


Figure 33: **IDM Core Metrics (T)**: Safety metrics are severely impacted in the unstable low-T regime due to deterministic jam nucleation.

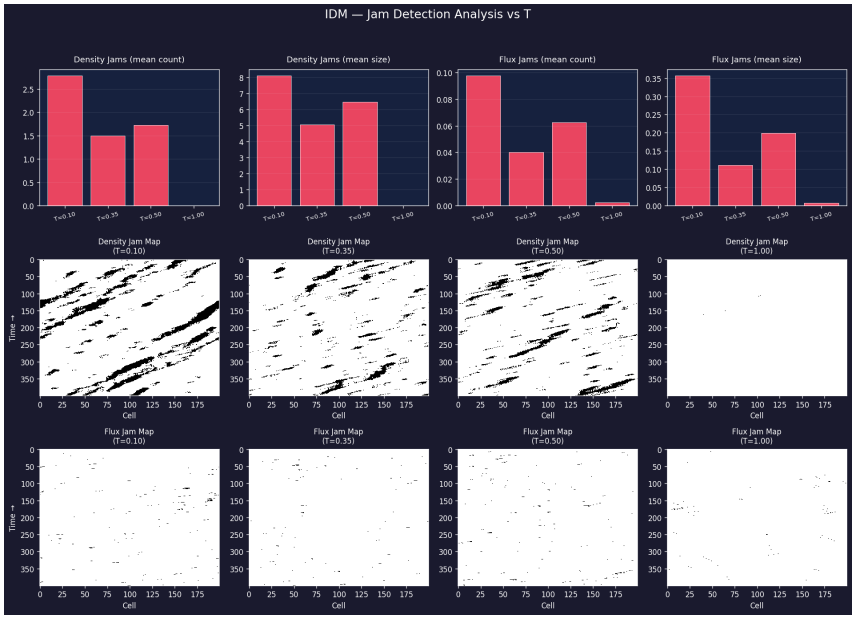


Figure 34: **IDM Jam Analysis (T)**: At low T, the system collapses into a single large jam; at high T, jams dissolve entirely.

Figure 35: **IDM Core Metrics** (ϵ): Near the stability boundary, even small noise triggers sharp degradation of safety and flow metrics.

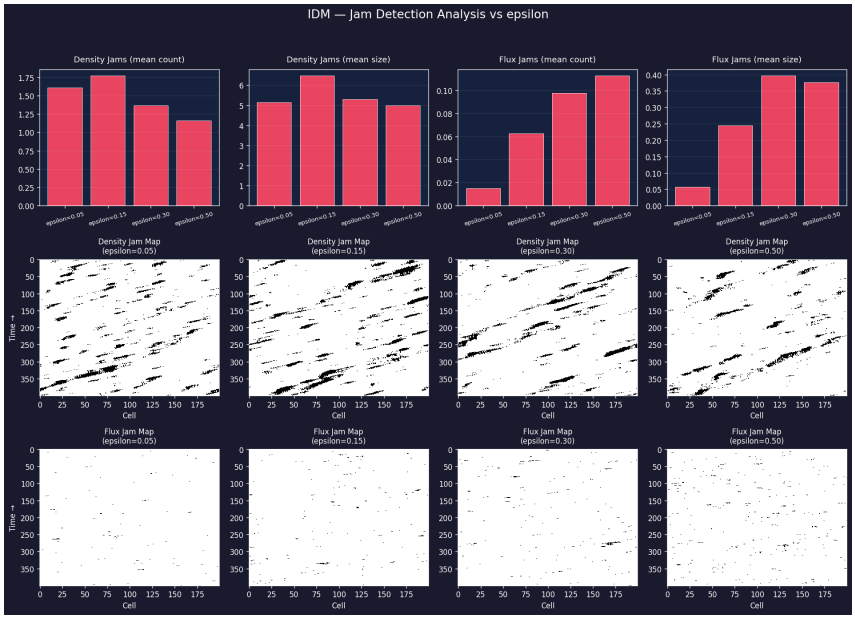


Figure 36: **IDM Jam Analysis** (ϵ): Noise-induced jams are spatially fragmented and short-lived compared to deterministic instabilities.

Impact of Kinematic Limits (a, $v_{\max}$, s_0)

Figures 37 and 39 explicitly visualize the throughput-versus-safety trade-off inherent to high a and $v_{\max}$ values. Figures 38 and 40 show these aggressive parameters result in sharper, shorter-lived jams. Finally, Figures **??** and 41 demonstrate how larger minimum spacing (s_0) effectively reduces gap-zero events but fundamentally lowers the internal density and capacity of congested regions.

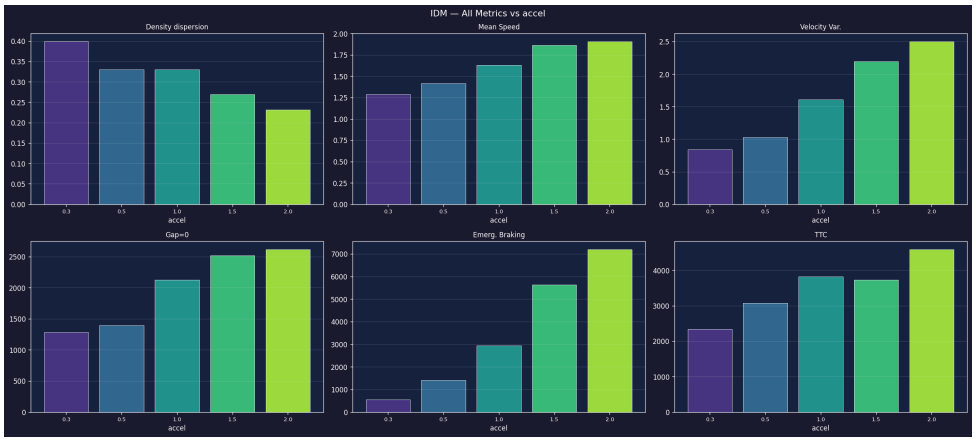


Figure 37: **IDM Core Metrics** (a): The trade-off between throughput and safety is visible in the divergence of mean velocity and TTC curves.

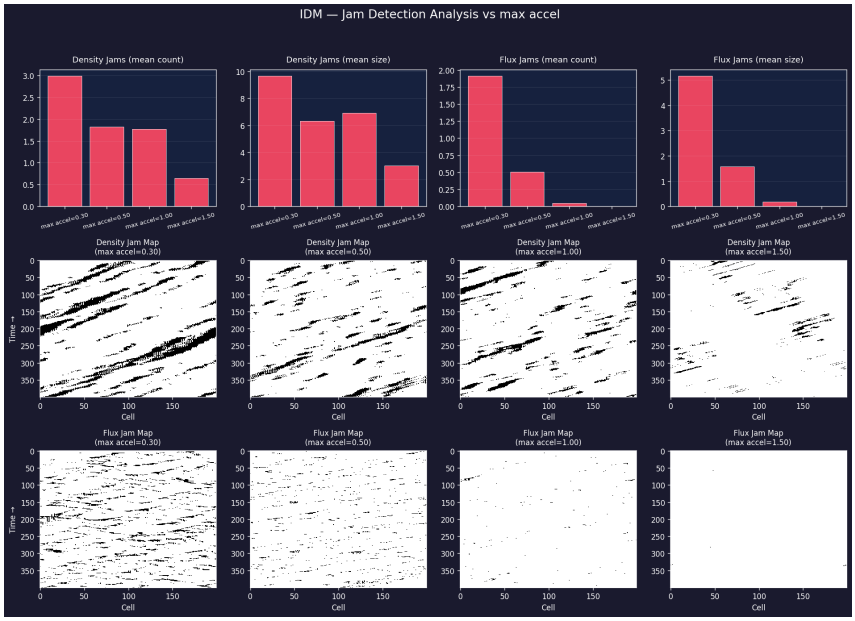


Figure 38: **IDM Jam Analysis** (a): Higher acceleration produces sharper, shorter-lived jams.

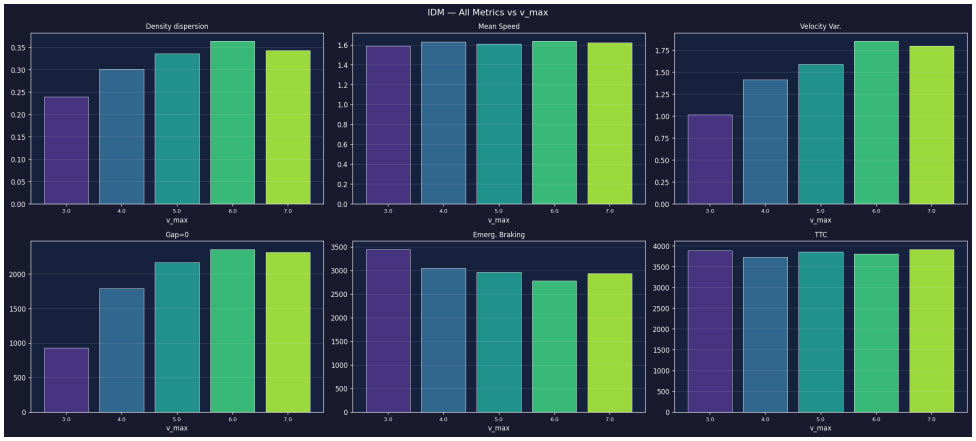


Figure 39: **IDM Core Metrics** ($v_{\max}$): Higher speed limits substantially degrade safety metrics by amplifying speed differentials at jam boundaries.

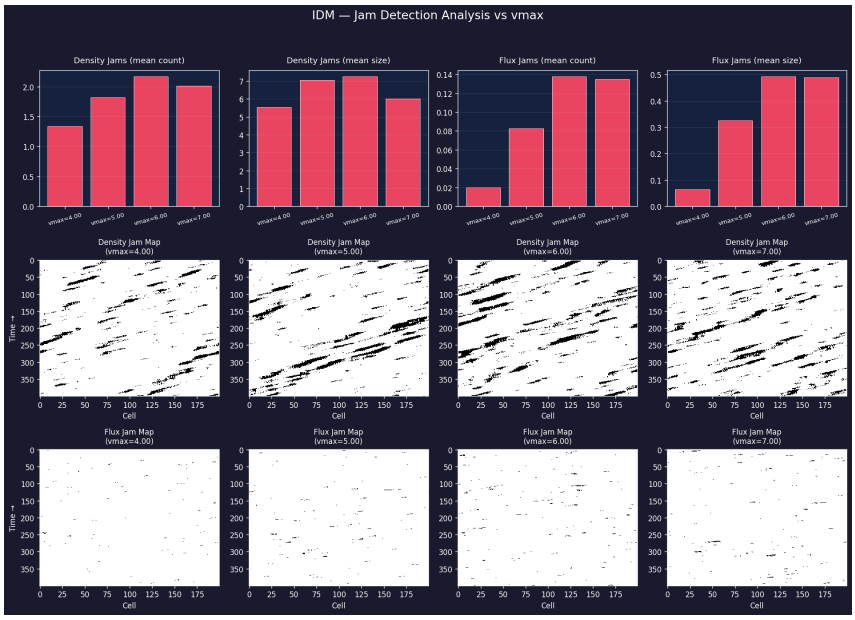


Figure 40: **IDM Jam Analysis** ($v_{\max}$): Flux-based detection captures sharp acceleration fronts more accurately than density-based mapping.

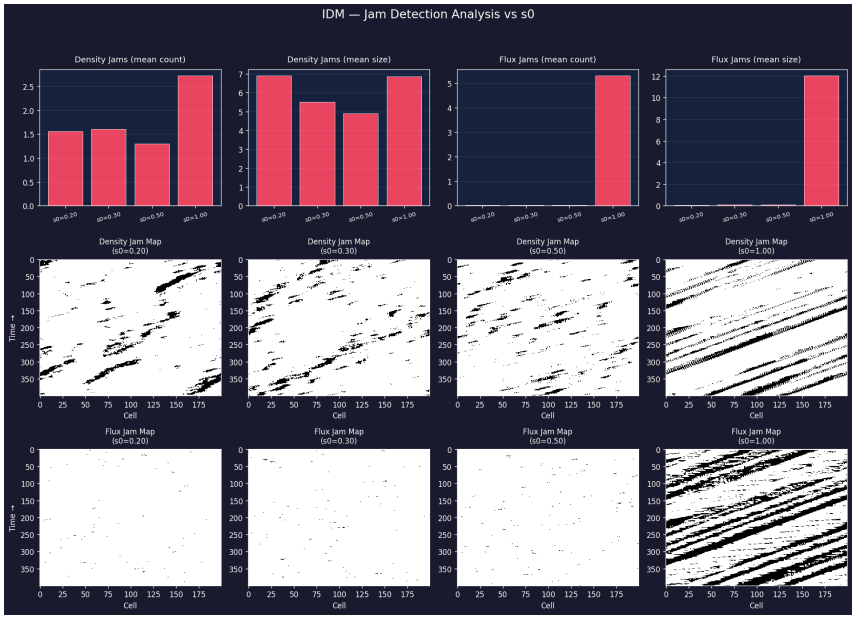


Figure 41: **IDM Jam Analysis** (s_0): Larger minimum spacing produces extended, low-density jams.

Implementation Notes

Code structure

The simulator is organised into four packages.

Package	Role
`environment/`	Ring road and graph topologies (`RingNetwork`, `GraphNetwork`)
`rules/`	Car-following models (`NaSchRule`, `KraussRule`, `IDMRule`)
`observers/`	Metric collectors (spacetime, flow, jam detection, safety)
`rl/`	DQN agent, training loop, evaluation utilities

The `Simulation` class wires them together: it holds a network and an update rule, steps the system synchronously, and notifies attached observers after each step.

Running a simulation

```
from environment.ring import RingNetwork
from rules.nasch import NaSchRule
from rules.idm import IDMRule
from rules.krauss import KraussRule
from simulation.simulation import Simulation

net  = RingNetwork(length=200, n_agents=40)
rule = NaSchRule(v_max=5, p_slow=0.3)      # or KraussRule / IDMRule
sim  = Simulation(net, rule, observers=[]) # attach Observer instances here
sim.run(n_steps=1000)                      # warm-up
sim.run(n_steps=5000)                      # measurement window
```

`IDMRule` and `KraussRule` share a similar constructor; their key parameters are listed in Section 3.

Notebooks

Notebook	Purpose
`simulation.ipynb`	Interactive NaSch animations and fundamental-diagram sweep.
`simulation_krauss_idm.ipynb`	Side-by-side Krauss and IDM space-time diagrams and fundamental curves.
`measure_congestion.ipynb`	Congestion detection: jam maps, slow-car tracking, flux and density observers.
`compare_micro_models.ipynb`	Head-to-head comparison of NaSch, Krauss, and IDM across densities.
`compare_baseline_vs_pslow0.ipynb`	Mixed-population sweep: fraction of $p_{\text{slow}} = 0$ deterministic cars vs. full NaSch baseline.
`train_rl_one_car.ipynb`	Step-by-step DQN assembly, training run configuration, and policy evaluation.
`compare_rl_n_agents_vs_baseline.ipynb`	Compare RL-controlled vs. pure baseline traffic across scenarios and numbers of controlled agents.
`radius_ablation_experiment.ipynb`	Train and evaluate RL policies with different observation radii.
`feature_ablation_experiment.ipynb`	Compare aggregate `mean` features versus ordered `atomic` neighbor features.

Command-line scripts

Train a DQN controller (NaSch background traffic).

```
python train_rl_one_car.py \
    --episodes 2000 --steps-per-episode 400 \
    --lr 3e-4 --gamma 0.97 \
    --save-path checkpoints/rl_one_car_dqn.pt
```

Saves a `.pt` checkpoint, a `.metrics.npz` training log, and a `.eval.json` summary.

Train ablation over observation radii (Krauss / IDM background).

```
python train_ablation.py --models krauss idm --radii=-1 1 2 3 --episodes 1500
```

Checkpoints are written to `checkpoints/ablation/rl_{model}_r{radius}.pt`.

Visualise RL vs baseline (single scenario).

```
python compare_rl_vs_baseline.py \
    --checkpoint checkpoints/rl_one_car_dqn.pt \
    --density 0.35 --p-slow 0.3 --steps 300
```

Aggregate ablation results.

```
python compare_ablation.py --models krauss idm --n-scenarios 24 --plot \
    --save-json results/ablation.json
```

Dependencies

`numpy`, `matplotlib`, `torch` (for RL), `scipy` (optional, DBSCAN jam clustering). Python 3.11+ is recommended.

F. Possible Extensions

Potential extensions not included in the main benchmark may include:

- brake-light or slow-to-start variants of Nagel–Schreckenberg;
- Gipps or FVDM car-following models;
- multilane roads and lane-changing rules;
- traffic lights and signal control;
- OSMnx-based extraction of a real road graph;
- Braess paradox experiments on a toy network;
- heterogeneous driver populations;
- game-theoretic route-choice dynamics.