2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

When Planners Meet Reality: How Learned, Reactive Traffic Agents Shift nuPlan Benchmarks

Steffen Hagedorn[1], Luka Donkov[2], Aron Distelzweig[3] and Alexandru P. Condurache[1]

***Abstract*— Planner evaluation in closed-loop simulation often uses rule-based traffic agents, whose simplistic and passive behavior can hide planner deficiencies and bias rankings. Widely used `IDM` agents simply follow a lead vehicle and cannot react to vehicles in adjacent lanes, hindering tests of complex interaction capabilities. We address this issue by integrating the state-of-the-art learned traffic agent model `SMART` into nuPlan. Thus, we are the first to evaluate planners under more realistic conditions and quantify how conclusions shift when narrowing the sim-to-real gap. Our analysis covers 14 recent planners and established baselines and shows that `IDM`-based simulation overestimates planning performance: nearly all scores deteriorate. In contrast, many planners interact better than previously assumed and even improve in multi-lane, interaction-heavy scenarios like lane changes or turns. Methods trained in closed-loop demonstrate the best and most stable driving performance. However, when reaching their limits in augmented edge-case scenarios, all learned planners degrade abruptly, whereas rule-based planners maintain reasonable basic behavior. Based on our results, we suggest `SMART`-reactive simulation as a new standard closed-loop benchmark in nuPlan and release the `SMART` agents as a drop-in alternative to `IDM` at https://github.com/shgd95/InteractiveClosedLoop.**

I. INTRODUCTION

In the development of automated driving systems it is crucial to assess performance under real-world conditions correctly before commencing tests in real traffic. The assessment's quality strongly depends on the evaluation procedure. Open-loop evaluation focuses on imitating a human expert driver as closely as possible. Various studies show that open-loop evaluation does not correlate with real-world driving performance [1], [2]. In contrast, closed-loop evaluation provides more generalizable results by simulating the environment's response to actions of the ego vehicle. Instead of imitating an expert driver, good driving behavior in general is rewarded in terms of safety, progress, and comfort metrics. Closed-loop simulation produces scene developments that deviate from the ground truth, for example the planner could decide to stay on a lane, whereas the expert driver performed a lane change in that situation, and achieve an equally good score. However, the generalizability of closed-loop evaluation strongly depends on the quality of the traffic agent model that updates the simulation, in which the planner acts. The widely used nuPlan framework and its benchmarks [3] rely on simplistic, rule-based `IDM`

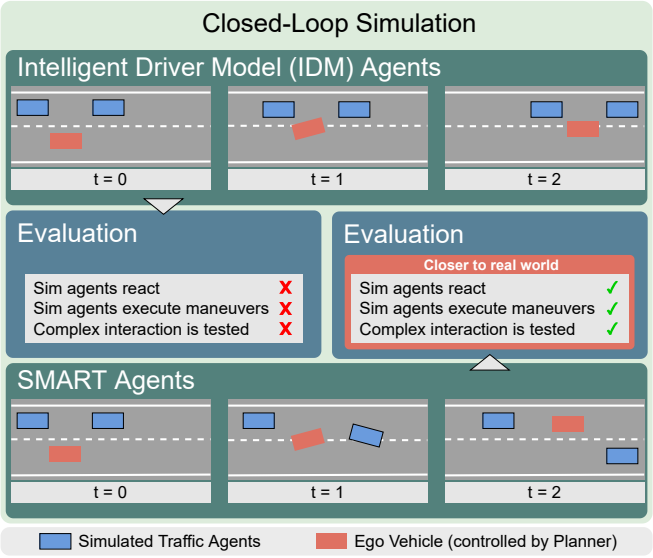

Fig. 1: Intelligent Driver Model (`IDM`) vs. `SMART` agents: same scene, different world. Starting with the same real-world scene state, the two traffic agent models lead to different scene developments when evaluating the same ego trajectory planner in a closed-loop simulation. Since the `IDM` cannot see vehicles in adjacent lanes it does not react to the ego vehicle's lane change attempt. In contrast, `SMART` agents trained from real traffic data behave less passive, react to other agents across lanes and execute driving maneuvers themselves. These characteristics of the sim agents are important for testing complex interactions and for assessing a planner's driving behavior in real-world traffic.

(Intelligent Driver Model) [4] traffic agents, whose behavior often lacks realism. For example, `IDM` agents do not perceive vehicles in adjacent lanes and ignore lane-change attempts (Fig. 1) or brake too hard when another car merges anyway. Such behavior is problematic for various reasons: Planners can aggressively take advantage of this passivity to exploit benchmarks [5]. Further, evaluation should rather cover the harder case in which other vehicles behave actively and human-like, instead of cautiously and passively. Finally, simplistic and unimodal traffic agent models prevent complex interactions and create a sim-to-real gap that biases planner benchmarks. We address this gap by integrating a learned, reactive state-of-the-art traffic agent model into the nuPlan framework to establish a new benchmark based on more realistic closed-loop simulation. We chose `SMART` [6] for its strong performance in the Waymo 2024 Sim Agents Challenge with high realism and interaction scores, map

[1]Robert Bosch GmbH, Leonberg, Germany and Institute for Neuro- and Bioinformatics, University of Lübeck, Germany `steffen.hagedorn@de.bosch.com`
[2]Robert Bosch GmbH, Feuerbach, Germany and Engineering Faculty, DHBW Stuttgart, Germany.
[3]Department of Computer Science, University of Freiburg, Germany.

compliance, and real-time inference, making it practical for large nuPlan rollouts.

Since changing the traffic agent model fundamentally alters closed-loop simulations, we perform a series of novel experiments to investigate how planners behave under more realistic conditions: Using the SMART agents we are the first ones to evaluate 14 state-of-the-art planners and common baselines of the nuPlan framework in a learned, reactive traffic simulation. We compare planner performance in SMART-based and IDM-based closed-loop simulations on three benchmarks: the comprehensive *Val14* [2], the more difficult *Test14-hard* [7], and the augmented edge-case scenarios of interPlan [8]. We find that IDM-based simulation indeed distorts planner rankings and systematically overestimates planner performance but underestimates interaction capabilities. Specifically, imitation-learned planners deteriorate in simple scenarios, while rule-based planners deteriorate in harder scenarios that require advanced interactions. Planners trained in closed-loop also perform better and more stable in realistic closed-loop simulations. When stress-testing planners in the hardest scenarios of interPlan, all methods reach their limits but rule-based methods degrade smoothly whereas learned planners exhibit a sudden tipping point.

In summary, the main contributions of our work are:

1) We are the first ones to evaluate established nuPlan planners in realistic, interactive traffic, revealing a general overestimation of planner performance under IDM-based simulation on diverse benchmarks alongside a simultaneous underestimation of interaction capabilities and further discrepancies.
2) We propose a new SMART-reactive benchmark for nuPlan to enable the realistic analysis of planner strengths and failure cases and introduce a corresponding closed-loop score.
3) We provide a community-ready implementation of state-of-the-art learned SMART traffic agents for nuPlan to enable model training and evaluation in realistic, reactive closed-loop simulations.

II. RELATED WORK

A. Automated Driving Evaluation

As it is unsafe and unethical to directly evaluate automated driving systems in public traffic, offline evaluation paradigms have emerged. Early planner evaluation has largely been *open-loop*: models predict the future based on fixed human-driven logs and are scored by imitation metrics with no feedback from the ego's actions to the scene. This setup scales well but suffers from covariate shift and does not reflect real driving quality [9], [1], [2].

In contrast, *closed-loop* evaluation restores interaction: the planner controls the ego in simulation and is judged by task success, safety, and comfort as the world reacts to its decisions (Fig. 2) [10]. Within closed-loop, benchmarks distinguish *non-reactive* backgrounds, which replay logged trajectories even if the ego deviates, from *reactive* backgrounds, where other agents respond to the ego. nuPlan

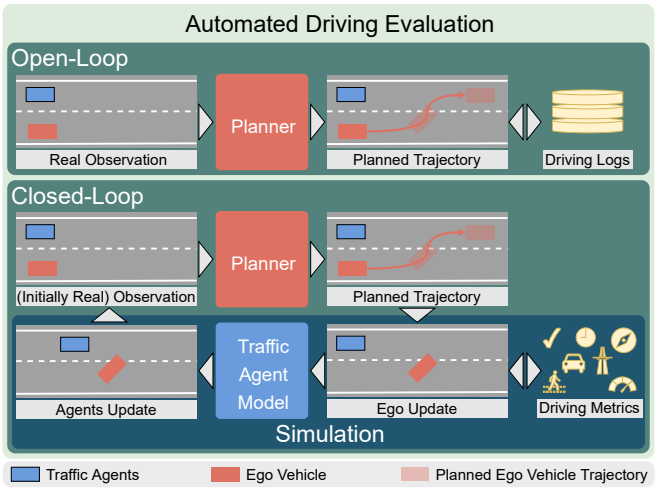

Fig. 2: Automated Driving Evaluation. In open-loop evaluation the planner generates a future trajectory from a real observation, which is then compared to an expert's driving log. However, open-loop performance does not correlate with real-world driving performance as it lacks evidence for the system's behavior when deviating from human expert states. Closed-loop evaluation addresses this shortcoming by allowing the scene state to deviate from the recorded real-world data and rewards good driving behavior instead of imitation. Nevertheless, the quality of the traffic agent model determines how well closed-loop performance mirrors real-world driving.

supports both and, by default, realizes reactivity with the rule-based Intelligent Driver Model (IDM) for background traffic [3], [4]. While simple and interpretable, such rule-based traffic can miss lateral negotiation, multimodality, or irrational (human-like) behavior, and thereby yield optimistic safety and overly stable planner rankings—precisely the gap we study by replacing IDM with learned, reactive agents (SMART) in nuPlan [6].

Multiple works show that conclusions shift when moving from non-reactive to reactive traffic: planner rankings, failure modes, and safety–efficiency trade-offs can flip [2], [11]. Swapping one reactive background for another also alters scene progress (Fig. 1), making the choice of the traffic agent simulation model crucial for the generalizability of evaluation results.

B. Traffic Agent Simulation

Closed-loop evaluation relies on a traffic simulator to model the scene progress. Traffic simulators can be classified as sensor-level or object-level methods. Sensor-level simulators render raw camera, lidar, or radar outputs, enabling full-stack and perception testing but with heavy compute costs. CARLA, LGSVL, and Bench2Drive are widely used simulators of this category [12], [13], [14]. Often sensor-level simulators use an underlying object-level simulation based on which the sensor outputs are rendered.

Object-level simulators update the world state without rendering sensor outputs. They are fast and scalable, making them well-suited for planning and interaction studies.

Their speed also enables rapid closed-loop training iterations. Perception robustness and full-stack performance, however, cannot be tested with object-level simulators. nuPlan, Nocturne, and WOSAC emphasize this paradigm with multi-agent, interaction-centric metrics and leaderboards [3], [15], [16]. Object-level simulators can be categorized into:

Rule-based. Classical simulation updates each vehicle individually with hand-crafted rules: `IDM` computes longitudinal acceleration for car-following and `MOBIL` decides when to change lanes [4], [17]. In practice, auxiliary rules for right-of-way, signals, and hand-tuned heuristics for gap acceptance and courtesy are added. The approach is simple, interpretable, and works well within a fixed domain. However, they struggle with lateral negotiation, multi-modal intent, and uncooperative behavior. They often react late and too strongly to cut-ins and need case-specific logic for traffic rules, signals, and every single maneuver, making broad generalization infeasible and limiting realism.

Hybrid. Hybrid models combine hand-crafted rules or heuristics with learned components. For example, `cogniBOT` adds learned components to a cognitive, rule-informed policy to improve realism while keeping control over simulation parameters [18]. Hybrid approaches aim to guarantee rule-compliance while gaining data-driven fidelity.

Learned. Earlier learned simulators such as `TrafficSim` showed that multi-agent traffic can be learned from logs and rolled out stably in closed loop [19]. This line of work paved the way for today's models. Recent research shows that training traffic agent models with reinforcement learning in closed-loop rollouts can improve purely imitation-based approaches [20]. Nevertheless, the recent trend focuses on imitation learning from logs: alongside generative approaches based on diffusion (`VBD` [21]) and variational autoencoders (`TrafficBots` [22]), the majority of models uses autoregressive Transformers (`Trajeglish` [23], `MVTE` [24], `SMART` [6]), `GUMP` [25], `BehaviorGPT` [26]).

Why we chose SMART. We integrate `SMART` as a new drop-in background for nuPlan simulations. `SMART` discretizes the vectorized scene and trajectories into spatio-temporal tokens and trains a decoder-only Transformer to predict the next tokens [6]. This design makes inference fast and memory-efficient, which is important for large `nuPlan` rollouts. `SMART` also achieved strong realism and interaction scores in the Waymo Sim Agents challenge, making it a good candidate to replace `IDM` when the goal is realistic evaluation of planners in reactive closed-loop simulation.

C. Trajectory Planning

While in traffic simulation the goal is to generate a realistic traffic flow, i.e., safe, traffic rule-compliant, and kinematically feasible diverse trajectories, the additional objective of trajectory planning is to move toward a navigation goal. Planner models span three broad styles, closely related to the design of the overall automated driving system in which they are applied [5].

End-to-end systems map sensor outputs like camera images or lidar point clouds directly to trajectories or controls. Sensor rendering is required to evaluate these models in closed-loop. For our experiments in the object-level simulation of nuPlan we exclude end-to-end systems and, instead, compare methods that are natively compatible with nuPlan's closed-loop simulation [3].

Modular systems keep a clear split between individual tasks, facilitating evaluation in object-level simulators [27]. They can be rule-based or learned. *Rule-based* approaches specify hand-crafted policies for car-following, lane-change, and search or optimization routines [4], [17], [2]. These are often used as a simple baseline for motion planning [4]. *Learned* planning modules are optimized on training data instead of relying on hand-crafted rules [28], [29], [30]. In the regime of learned planners, two training paradigms dominate. *Imitation learning (IL)* fits policies to human demonstrations and remains the most frequently applied paradigm because of data scale and training stability [29], [31]. IL can achieve strong closed-loop scores but is sensitive to covariate shift through accumulating errors, leading to states unseen during training without closed-loop feedback [2], [1]. To address this issue, *Reinforcement Learning (RL)* trains on closed-loop rollouts to let the model deviate from logged expert states and explore the consequences of its actions [20]. Instead of fitting the policy to human demonstrations, RL optimizes a reward term that scores generally desirable driving behavior like collision avoidance, staying on-road, or traffic rule compliance [32].

Hybrid planners combine learned components with hand-crafted rules. One typical combination is to generate an initial plan with a learned model, that is then fine-tuned by explicit optimization algorithms [33], [34]. Other methods use rule-based safety layers to enforce drivable area compliance, collision avoidance and traffic rule-adherence [2].

For our study we select diverse trajectory planners that cover rule-based and hybrid methods as well as IL-based and RL-based learned modular approaches to investigate how realistic traffic agent simulation affects their performance and if conclusions related to these paradigms can be drawn.

III. METHODOLOGY

A. Datasets & Benchmarks

We train the `SMART` model and conduct all planner studies on the nuPlan dataset and benchmark [3], which contains $\sim$1,300 hours of expert driving with auto-labeled tracks, traffic lights, and scenario tags. nuPlan's closed-loop simulator runs each scenario for 15 s at 10 Hz in which the ego vehicle is controlled by the candidate planner and a low-level controller. Background traffic can be run in two modes: *non-reactive* (log replay) and *reactive* (agents controlled by the `IDM`). We adopt the *train_150k* training split that samples 150,000 scenarios across all scenario types, and evaluate on the *Val14* benchmark, which includes up to 100 scenarios per scenario type across 14 types, yielding 1,090 total scenarios after excluding a small number of reactive runs that fail to initialize. (This matches the setup used by recent nuPlan papers.) For additional experiments we use *Test14-hard*, a curated 280-scenario split formed by running

100 candidates per scenario type with the strong rule-based baseline PDM-Closed and selecting the 20 lowest-scoring, i.e., 'hardest' cases, emphasizing long-tail interactive failures like in tight merges and unprotected turns [7]. To stress-test planners under even harder conditions we also perform experiments on the manually augmented *interPlan* benchmark [8]. Specifically, we use 30 lane change scenarios of which ten are in low, medium and high-density traffic, each.

To assess the realism of SMART simulation agents beyond nuPlan, we additionally curate results of the the Waymo Open Sim Agents Challenge (WOSAC) 2024, that is based on the Waymo Open Motion Dataset (WOMD) [16]. WOMD provides large-scale, object-level trajectories with HD maps, comprising over 100k 20 s segments at 10 Hz.

B. Metrics

We report metrics for two aspects of our study: (i) planner performance and (ii) the realism of simulated traffic agents.

(i) Planner evaluation. nuPlan employs three principal metrics: the *open-loop score (OLS)*, the *non-reactive closed-loop score (CLS-NR)*, and the *reactive closed-loop score (CLS-R)* [3]. Consistent with prior findings that open-loop prediction quality weakly correlates with closed-loop driving effectiveness, we focus on closed-loop performance and provide the OLS only for reference. Closed-loop scores are computed per scenario as a weighted combination of soft metrics: *progress* along the route, *time-to-collision* within bounds, and *comfort* (e.g., lateral/longitudinal jerk), which are subject to hard multipliers that zero the score if *at-fault collisions* or *drivable-area violations* occur [35]. Scenario scores are then averaged across the benchmark into a normalized $[0, 100]$ composite used for ranking [3], [2].

SMART-reactive score. To narrow the sim-to-real gap of closed-loop planner evaluation, we propose a new benchmark on the *Val14* and *Test14-hard* dataset splits of nuPlan. Our benchmark uses the scenarios of *Val14* and *Test14-hard* and introduces a novel *SMART-reactive closed-loop score (CLS-SR)*. The score is computed from the same metrics as CLS-R but utilizes SMART to update the simulated vehicles instead of IDM. This allows direct comparison of planners across non-reactive, rule-based reactive, and learned reactive simulation.

(ii) Simulated-traffic realism. For SMART and baselines, we report *Average Displacement Error (ADE)* compared to the logged expert driver ground truth and the *Realism Meta-Metric (RMM)* as defined by WOSAC [16]. The RMM rewards realistic traffic flow instead of imitation by aggregating kinematic-based, map-compliance, and interaction-oriented components into a single score [16].

C. SMART Integration

Training. We transfer and train SMART in the nuPlan framework while avoiding architectural changes compared to the published model [6]. Concretely, we use the 2 Hz action-token vocabulary with 1024 discrete road and motion tokens, each. We implement the preprocessing of nuPlan agent and map data to match the preprocessing on WOMD. Following SMART, we introduce noise in the tokenization pipeline: when discretizing trajectories, we perturb the currently matched token by randomly selecting one of the top-6 nearest tokens to the ground-truth in the vocabulary and continue matching from the perturbed state, which exposes the model to distribution shift and mitigates compounding-errors. The generated input format is directly compatible with SMART and facilitates training without any adaptions of the model itself. Our models are trained with a 1 s history and an 8 s ground-truth future. Training follows teacher forcing with cross-entropy for next-token prediction. We trained the small SMART model with 7M learnable parameters on varying amounts of nuPlan and WOMD data to convergence and compared performance on a separate validation set. We selected checkpoints by token accuracy. The model used in all subsequent experiments, and released with this paper, was trained on the nuPlan *train_150k* split for eight epochs. On a single NVIDIA H200 GPU with a mini-batch size of four, training required four days.

Inference & runtime. nuPlan advances the simulation at 10 Hz. SMART operates on 2 Hz tokens, so we integrate it in a receding-horizon fashion: every 0.5 s we re-encode the current scene (including the last 1 s of history), run a single decoding step to forecast agent movements, and select the maximum-probability token to make the simulation deterministic. We then upsample the 2 Hz tokens to 10 Hz trajectories using SMART's upsampling procedure and let nuPlan's tracker execute that motion for the next 0.5 s. This avoids rolling out a full 8 s token sequence at once, keeps computation bounded, and ensures tight feedback between planner actions and background reactions. The integration is packaged as a drop-in reactive background for nuPlan, so users can select SMART in the simulator configuration exactly like the standard IDM agents.

D. Planner Models

We benchmark a diverse set of nuPlan-compatible planners that have public code and released checkpoints, so that results are not confounded by re-training variance. The selection spans rule-based, hybrid, imitation-learning, and reinforcement-learning approaches, including influential earlier baselines and recent state of the art.

- IDM [4] implements rule-based car-following and serves as a conservative planning baseline provided in nuPlan.
- PDM-Closed [2] is a rule-based, centerline-following planner that assesses and optimizes closed-loop metrics.
- PDM-Hybrid [2] combines a learned ego-forecast with a rule-based refinement stage to improve stability.
- GameFormer [36] combines a Transformer with game theory to model strategic interactions among traffic participants, paired with a rule-based refinement.
- DTPP [34] stands for differentiable tree-structured policy planning that jointly learns prediction and planning costs and searches over a trajectory tree.
- PLUTO [31] implements modular end-to-end imitation learning with vectorized scene encoding and contrastive training, leading to high closed-loop performance.

- UrbanDriver [28] is an early learned baseline integrated in nuPlan, that realizes a policy-gradient planner trained from demonstrations in object-level closed loop.
- GC-PGP [37] extends naïve behavior cloning with goal-conditioning at inference time.
- PlanCNN [29] realizes a rasterized BEV planner that scores candidate trajectories with a convolutional backbone and chooses the best.
- PlanTF [7] is an early Transformer-based planner operating on vectorized map and agent features.
- PDM-Open [2] is a fully learned, simplistic variant of the PDM conditioned on a reference path and ego state.
- Diffusion Planner [30] applies a generative model that samples ego trajectories via diffusion while enforcing map and interaction constraints.
- CaRL [32] implements a large-scale reinforcement-learning planner optimized directly in closed-loop training with simple rewards.

To enable fair evaluation in nuPlan's closed-loop simulator, we use each author's recommended checkpoint and default evaluation configuration.

IV. EXPERIMENTS & RESULTS

Running the first ever closed-loop planner evaluation with realistically reactive traffic agents on nuPlan yields the results shown in Table I. The comparison comprises 14 diverse planners and reports our novel SMART-reactive closed-loop score (CLS-SR), alongside the standard IDM-reactive score (CLS-R) and the non-reactive log-replay score (CLS-NR). Open-loop scores are provided for reference; the focus of our analysis is on closed-loop simulations. All closed-loop runs were executed by us under identical simulator settings to ensure comparability. We structure the analysis around key insights emerging from these results.

Imitation-learned planners deteriorate on simple scenarios, rule-based planners on hard ones. Table I shows that, on the standard *Val14* benchmark, imitation-learned planners experience the most considerable drop when moving from IDM agents to SMART agents in closed-loop evaluation (difference of CLS-SR − CLS-R). On average, imitation-learned planners have a -5.17 decreased closed-loop score, whereas hybrid models only decrease by -3.25 and rule-based planners by just -2.0. This highlights the advantage of injecting rule-based knowledge to guarantee collision avoidance and drivable-area compliance and to bound comfort-related quantities such as jerk. An exception among the learned methods is PDM-Open, which remains nearly constant across backgrounds. We attribute this to its minimalistic inputs (ego state and centerline), which effectively encode a centerline-following prior: while unsuited to maximize the absolute score it yields stable behavior.

Test14-hard addresses the imbalance of *Val14* toward easy scenarios by selecting the 20% lowest-scoring cases per type under a strong baseline (PDM-Closed). Repeating our experiments on *Test14-hard* reverses the trend: imitation-learned planners lose less performance (-2.83) than hybrid (-4.75) and rule-based (-4.5) methods. Hard scenarios require nuanced interaction skills like gap negotiation, multi-lane merges, unprotected turns, that fixed rules struggle to express. Our conclusion is that rule-based structure stabilizes behavior on simple to medium difficulty, but does not generalize as well to challenging interaction-heavy scenes. Nevertheless, all purely imitation-learned planners are still outperformed by rule-based and hybrids methods when looking at the absolute CLS-SR.

Closed-loop training enables stable high-quality performance. As outlined above, even the state-of-the-art IL model Diffusion Planner cannot match the performance of the rule-based PDM-Closed. Our experiments with the RL model CaRL confirm that this finding is related to the covariate shift between open-loop training and closed-loop application of the selected IL models (Table I). CaRL is the only available RL model that is trained on closed-loop rollouts of the nuPlan simulator. While it already matched the performance of PDM-Closed on Val14-CLS-R it ranks first on our SMART-based *Val14*-CLS-SR benchmark and clearly stands out against all other models on *Test14-hard*-CLS-SR with a +8 margin on the second-placed PDM-Closed. The performance is also more stable than that of many IL, hybrid, and rule-based planners when switching from IDM to SMART-based simulation. We relate these results to (i) the property of reinforcement learning that rule-based knowledge can be injected via reward shaping while still learning a flexible policy from large-scale datasets and especially (ii) the exposure of the policy to its actions in closed-loop rollouts during training, effectively addressing the covariate-shift problem. Our experiments show that after years of PDM-Closed leading nuPlan benchmarks, CaRL finally clearly outperforms the rule-based method.

IDM agents distort planning benchmarks. To understand why performance deteriorates when switching from an IDM background to SMART, we decompose the composite CLS into its components and break down results by scenario type on *Val14* (Fig. 4, left). Learned planners benefit disproportionately from IDM's passive and cautious behavior: time-to-collision, ego progress, comfort, and at-fault collisions degrade most under SMART. This quantifies that inflated planner scores can be facilitated by IDM creating large gaps and yielding early. In contrast, drivable-area compliance and speed-limit compliance metrics, that are less directly tied to the behavior of other drivers, remain comparatively stable.

The scenario-wise analysis (Fig. 4, right) reveals a consistent pattern. In multi-lane, interaction-heavy settings like lane change, starting right/left turn, starting straight traffic-light intersection traversal, and traversing pickup/dropoff, scores decrease less or even increase under SMART. The reason is that IDM primarily reacts to the lead vehicle in-lane, whereas SMART attends to all nearby traffic, enabling realistic cross-lane interactions. The extreme case occurs when an IDM planner is evaluated against IDM agents: neither side perceives adjacent-lane vehicles, leading to collisions and abrupt driving actions when new vehicles suddenly 'appear' in the lane. Switching to SMART agents can raise the same planner's CLS by up to 12%. In summary,

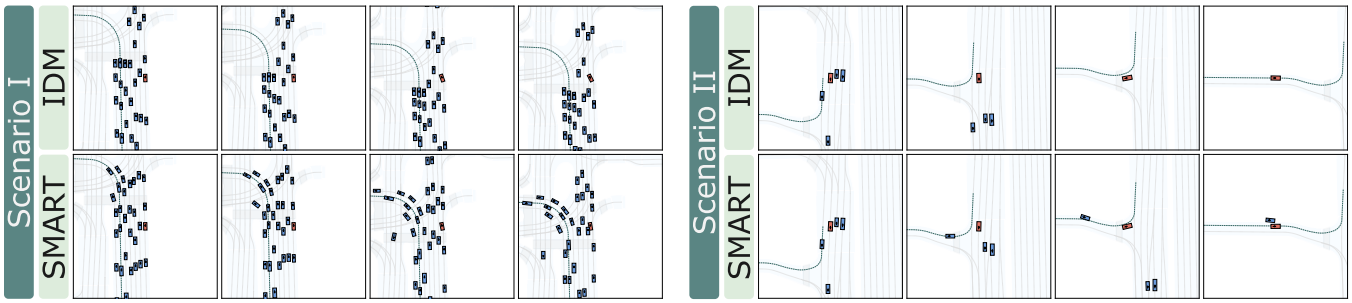

Fig. 3: Closed-loop simulation with SMART agents compared to IDM agents in two exemplary scenarios of interPlan-lane-change. In scenario I the ego (red) must execute multiple lane changes in dense traffic to follow the intended route (dashed). SMART agents show more diverse behavior and decreased passivity compared to IDM. The ego surprises the other vehicle with its sudden lane change inside the intersection, causing a collision. In scenario II a SMART agent turns right and then stops at the side of the road. The ego vehicle nudges around it after turning, proving its interaction capability, which remains untested when using IDM agents. In all four simulations the ego vehicle is steered by the best performing planner CaRL.

Planner			Val14				Test14-hard	
Type	**Paradigm**	**Method**	**OLS ↑**	**CLS-NR ↑**	**CLS-R ↑**	**CLS-SR ↑**	**CLS-R ↑**	**CLS-SR ↑**
Expert Log	Human	Log Replay [3]	100	94	80	76 (-4)	69	64 (-5)
Rule-Based	Rule	IDM Planner [4]	38	76	77	77 (0)	62	55 (-7)
Rule-Based	Rule	PDM-Closed [2]	42	93	93	89 (-4)	76	74 (-2)
Hybrid	IL + Rule	PDM-Hybrid [2]	84	93	93	89 (-4)	76	72 (-4)
Hybrid	IL + Rule	GameFormer [36]	82	81	82	78 (-4)	70	62 (-8)
Hybrid	IL + Rule	DTPP [34]	61	64	64	62 (-2)	47	47 (0)
Hybrid	IL + Rule	PLUTO [31]	—	93	87	84 (-3)	76	69 (-7)
Learned	IL	Urban Driver [28]	82	53	51	43 (-8)	41	38 (-3)
Learned	IL	GC-PGP [37]	83	59	56	50 (-6)	43	41 (-2)
Learned	IL	PlanCNN [29]	64	73	70	65 (-5)	58	51 (-7)
Learned	IL	PlanTF [7]	89	85	77	72 (-5)	59	58 (-1)
Learned	IL	PDM-Open [2]	86	50	55	53 (-2)	37	38 (+1)
Learned	IL	Diffusion Planner [30]	—	90	83	78 (-5)	68	63 (-5)
Learned	RL	CaRL [32]	—	94	93	90 (-3)	85	82 (-3)

TABLE I: **Planner evaluation on nuPlan benchmarks.** Open-loop (OLS) and closed-loop non-reactive (CLS-NR) results are reported according to [2], [29], and [30], while all closed-loop reactive (CLS-R) and closed-loop SMART reactive (CLS-SR) scores are from our experiments. CLS-SR is measured using our SMART sim agents.

many planners interact better than their IDM-based scores suggest. Low scores often rather reflect deficiencies of the IDM background.

Beyond their limits learned planners degrade abruptly. Finally, we stress-test planners on augmented interPlan lane-change scenarios that vary traffic density (Table II). interPlan increases density by adding agents to real-world scenarios while preserving map and route structure. At low density, lane changes remain feasible, and several planners improve under SMART due to more cooperative across-lane reactions. As density increases, absolute performance drops sharply for nearly all methods: gaps close, negotiations become contested, and safe merges disappear. SMART benefits only the most recent learned methods (PLUTO, Diffusion Planner, CaRL) at mid density, but most planners deteriorate as interaction complexity and contention rise. At high density, lane changes are often infeasible: SMART agents still behave realistically, but the task becomes near-impossible (Fig. 3, Scenario I). Rule-based and hybrid policies fail more gracefully (e.g., declining to change lanes), whereas fully learned methods degrade more abruptly—consistent with being pushed far out-of-distribution. This abrupt degradation highlights the critical need for a supervisory rule-based safety layer to guarantee a minimal level of safe performance. Notably, CaRL, that is trained with IDM agents, clearly loses performance at high density when switching to SMART agents, underscoring sensitivity to the training background and the need for diverse, realistic agents during training. We suggest that using SMART agents not only for evaluation but already in closed-loop training could advance planning further.

SMART agents narrow the sim-to-real gap. All experiments rest on the premise that a learned, reactive simulator such as SMART more accurately reproduces real traffic than rule-based IDM. We therefore compare these two models directly. Fig. 3 shows a qualitative comparison of traffic flow when using the two models for traffic agent simulation. SMART agents demonstrate more versatile behavior than IDM

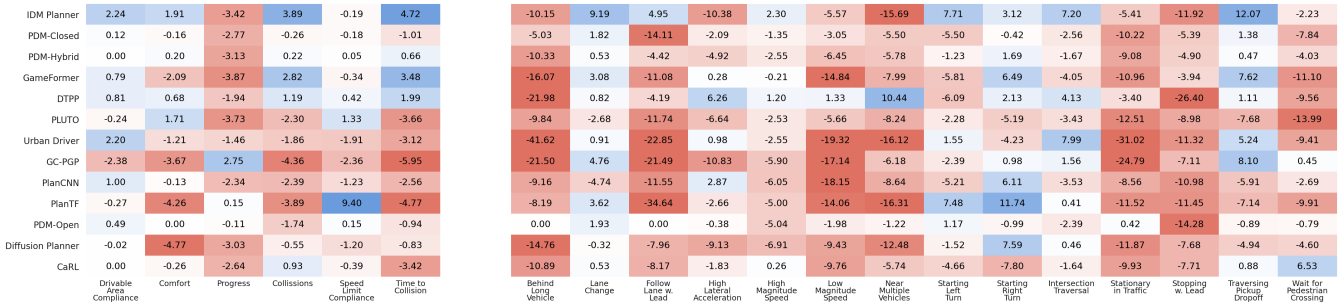

Fig. 4: Planner performance shift under SMART agents compared to IDM agents on *Val14* (CLS-SR − CLS-R). The left plot shows a metric-wise breakdown per planner. The right plot provides a scenario-wise analysis.

	CLS ↑					
	Low Density		Mid Density		High Density	
Planner	IDM	SMART	IDM	SMART	IDM	SMART
IDM	63	59 (-4)	62	58 (-4)	63	53 (-10)
PDM-Closed	61	62 (+1)	62	54 (-8)	62	46 (-16)
PDM-Hybrid	62	62 (0)	62	55 (-7)	61	46 (-15)
GameFormer	16	30 (+14)	47	47 (0)	28	37 (+9)
DTPP	65	58 (-7)	66	56 (-10)	73	24 (-49)
PLUTO	67	66 (-1)	42	43 (+1)	48	37 (-11)
Urban Driver	0	0 (0)	29	15 (-14)	0	0 (0)
GC-PGP	27	35 (+8)	17	10 (-7)	61	0 (-61)
PlanCNN	46	56 (+10)	68	12 (-56)	61	0 (-61)
PlanTF	50	34 (-16)	41	37 (-4)	74	42 (-32)
PDM-Open	23	33 (+10)	21	14 (-7)	26	0 (-26)
Diff. Planner	40	21 (-19)	21	40 (+19)	16	27 (+11)
CaRL	63	69 (+6)	38	49 (+11)	66	29 (-37)

TABLE II: **Closed-loop scores on the interPlan [8] lane change benchmark.** The values reported for SMART agents correspond to CLS-SR and those for IDM to CLS-R. Rule-based planners degrade smoothly while most methods with learned components hit sudden tipping points when inferred too far outside of the training distribution (mid/high density).

	nuPlan Val14	WOSAC 2024			
Method	**ADE@8s ↓**	**RMM ↑**	**Kinematic ↑**	**Interactive ↑**	**Map ↑**
IDM [4]	9.60	0.62	0.48	0.72	0.56
SMART 7M [6]	0.75	0.76	0.48	0.80	0.86

TABLE III: **Sim agent performance across nuPlan and WOMD benchmarks.** SMART provides indeed more realistic agents than IDM. nuPlan IDM results are based on [38], WOMD results are curated from [6], [39].

agents, including U-turns, lane changes, and even stopping at the road side. Especially the SMART agents on the turning lanes are less passive and enter the intersection much sooner.

To quantify these observations, we compiled published metrics where available and reproduced any missing quantities under the same protocol. First, we consider ADE over 8 s to evaluate the imitation quality. As summarized in Table III, the SMART agents clearly imitate real-world traffic logs more accurately (0.75 m) than the rule-based IDM agents (9.6 m).

However, pure imitation accuracy is an imperfect measure for simulation quality, since future trajectories are inherently multi-modal and many distinct futures can be acceptable. We therefore complement ADE with the realism meta metric, a realism-oriented composite that evaluates agent behavior along three axes: kinematics, interactions, and map compliance. Table III shows that SMART is on par with IDM for kinematic consistency, exhibits stronger interactions between agents, and considerably improves map compliance, reducing off-road events and maintaining larger, human-like distances to road edges. While SMART is not flawless, these results support its use as a more realistic reactive background than IDM. Evaluating planners under SMART therefore narrows the sim-to-real gap and enables a more faithful assessment of planning behavior.

V. CONCLUSION AND FUTURE WORK

Current nuPlan closed-loop evaluations rely on simple, rule-based reactive traffic (IDM), that can bias results and distort rankings. We addressed this by integrating learned reactive SMART agents into nuPlan and introducing a more realistic benchmark for nuPlan. On the new benchmark we evaluated 14 established planners and compared them to their scores in IDM simulation. Our experiments confirm that IDM shifts benchmarks and is often the origin of planners' poor interaction results, whereas SMART narrows the sim-to-real gap. We further find that: (i) imitation-learned planners tend to deteriorate on simple scenarios while rule-based planners deteriorate on hard, interaction-heavy ones; (ii) closed-loop training yields more stable, high-quality driving; and (iii) when pushed beyond their limits, learned planners degrade abruptly. Overall, the reinforcement-learned CaRL performs best in our simulations and clearly surpasses long-time baseline PDM-Closed on *Test14-hard*. By releasing drop-in SMART agents for nuPlan and proposing standardized reporting (CLS-SR), we enable the community to compare planners under traffic that responds credibly to the ego. Promising next steps include training CaRL directly against SMART agents and probing robustness with non-deterministic agent behavior, which our implementation supports.

ACKNOWLEDGMENT

We thank the authors of SMART for providing their codebase, on which we built our work, and for the helpful exchanges that saved us from running into experimental dead ends. We also thank our colleague Bin Yang for facilitating our communication with the SMART authors.

REFERENCES

[1] F. Codevilla, A. M. Lopez, V. Koltun, and A. Dosovitskiy, "On offline evaluation of vision-based driving models," in *Proceedings of the European conference on computer vision*, 2018, pp. 236–251.

[2] D. Dauner, M. Hallgarten, A. Geiger, and K. Chitta, "Parting with misconceptions about learning-based vehicle motion planning," in *Conference on Robot Learning*. PMLR, 2023, pp. 1268–1281.

[3] N. Karnchanachari, D. Geromichalos, K. S. Tan, N. Li, C. Eriksen, S. Yaghoubi, N. Mehdipour, G. Bernasconi, W. K. Fong, Y. Guo *et al.*, "Towards learning-based planning: The nuplan benchmark for real-world autonomous driving," in *2024 IEEE International Conference on Robotics and Automation*, pp. 629–636.

[4] M. Treiber, A. Hennecke, and D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Physical review E*, vol. 62, no. 2, p. 1805, 2000.

[5] S. Hagedorn, M. Hallgarten, M. Stoll, and A. P. Condurache, "The integration of prediction and planning in deep learning automated driving systems: A review," *IEEE Transactions on Intelligent Vehicles*, 2024.

[6] W. Wu, X. Feng, Z. Gao, and Y. Kan, "Smart: Scalable multi-agent real-time motion generation via next-token prediction," *Advances in Neural Information Processing Systems*, vol. 37, pp. 114 048–114 071, 2024.

[7] J. Cheng, Y. Chen, X. Mei, B. Yang, B. Li, and M. Liu, "Rethinking imitation-based planners for autonomous driving," in *2024 IEEE International Conference on Robotics and Automation*, pp. 14 123–14 130.

[8] M. Hallgarten, J. Zapata, M. Stoll, K. Renz, and A. Zell, "Can vehicle motion planning generalize to realistic long-tail scenarios?" in *2024 IEEE International Conference on Intelligent Robots and Systems*, pp. 5388–5395.

[9] S. Ross, G. Gordon, and D. Bagnell, "A reduction of imitation learning and structured prediction to no-regret online learning," in *Proceedings of the fourteenth international conference on artificial intelligence and statistics*. JMLR Proceedings, 2011, pp. 627–635.

[10] M.-K. Bouzidi, C. Schlauch, N. Scheuerer, Y. Yao, N. Klein, D. Göhring, and J. Reichardt, "Closing the loop: Motion prediction models beyond open-loop benchmarks," *arXiv preprint arXiv:2505.05638*, 2025.

[11] Y. Fan, Y. Li, and S. Wang, "Risk-aware self-consistent imitation learning for trajectory planning in autonomous driving," in *European Conference on Computer Vision*. Springer, 2024, pp. 270–287.

[12] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "Carla: An open urban driving simulator," in *Conference on robot learning*. PMLR, 2017, pp. 1–16.

[13] G. Rong, B. H. Shin, H. Tabatabaee, Q. Lu, S. Lemke, M. Možeiko, E. Boise, G. Uhm, M. Gerow, S. Mehta *et al.*, "Lgsvl simulator: A high fidelity simulator for autonomous driving," in *2020 IEEE 23rd International conference on intelligent transportation systems*, pp. 1–6.

[14] X. Jia, Z. Yang, Q. Li, Z. Zhang, and J. Yan, "Bench2drive: Towards multi-ability benchmarking of closed-loop end-to-end autonomous driving," *Advances in Neural Information Processing Systems*, vol. 37, pp. 819–844, 2024.

[15] E. Vinitsky, N. Lichtlé, X. Yang, B. Amos, and J. Foerster, "Nocturne: a scalable driving benchmark for bringing multi-agent learning one step closer to the real world," *Advances in Neural Information Processing Systems*, vol. 35, pp. 3962–3974, 2022.

[16] N. Montali, J. Lambert, P. Mougin, A. Kuefler, N. Rhinehart, M. Li, C. Gulino, T. Emrich, Z. Yang, S. Whiteson *et al.*, "The waymo open sim agents challenge," *Advances in Neural Information Processing Systems*, vol. 36, pp. 59 151–59 171, 2023.

[17] A. Kesting, M. Treiber, and D. Helbing, "General lane-changing model mobil for car-following models," *Transportation Research Record*, vol. 1999, no. 1, pp. 86–94, 2007.

[18] L. Brostek, C. Rössert, J. Drever, and A. Knorr, "Achieving realism in traffic simulations: Performance of a cognitive behavior model on the waymo open sim agent challenge," cogniBIT GmbH, Technical Report, 2024. [Online]. Available: https://www.cognibit.ai/docs/Achieving_Realism_in_Traffic_Simulations-whitepaper.pdf

[19] S. Suo, S. Regalado, S. Casas, and R. Urtasun, "Trafficsim: Learning to simulate realistic multi-agent behaviors," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2021, pp. 10 400–10 409.

[20] M. Bitzer, R. Cimurs, B. Coors, J. Goth, S. Ziesche, P. Geiger, and M. Naumann, "Analyzing closed-loop training techniques for realistic traffic agent models in autonomous highway driving simulations," *arXiv preprint arXiv:2410.15987*, 2024.

[21] Z. Huang, Z. Zhang, A. Vaidya, Y. Chen, C. Lv, and J. F. Fisac, "Versatile behavior diffusion for generalized traffic agent simulation," *arXiv preprint arXiv:2404.02524*, 2024.

[22] Z. Zhang, C. Sakaridis, and L. Van Gool, "Trafficbots v1. 5: Traffic simulation via conditional vaes and transformers with relative pose encoding," *arXiv preprint arXiv:2406.10898*, 2024.

[23] J. Philion, X. B. Peng, and S. Fidler, "Trajeglish: Traffic modeling as next-token prediction," *arXiv preprint arXiv:2312.04535*, 2023.

[24] Y. Wang, T. Zhao, and F. Yi, "Multiverse transformer: 1st place solution for waymo open sim agents challenge 2023," *arXiv preprint arXiv:2306.11868*, 2023.

[25] Y. Hu, S. Chai, Z. Yang, J. Qian, K. Li, W. Shao, H. Zhang, W. Xu, and Q. Liu, "Solving motion planning tasks with a scalable generative model," in *European Conference on Computer Vision*. Springer, 2024, pp. 386–404.

[26] Z. Zhou, H. Haibo, X. Chen, J. Wang, N. Guan, K. Wu, Y.-H. Li, Y.-K. Huang, and C. J. Xue, "Behaviorgpt: Smart agent simulation for autonomous driving with next-patch prediction," *Advances in Neural Information Processing Systems*, vol. 37, pp. 79 597–79 617, 2024.

[27] S. Hagedorn, A. Distelzweig, M. Hallgarten, and A. P. Condurache, "Learning through retrospection: Improving trajectory prediction for automated driving with error feedback," in *2025 IEEE International Conference on Intelligent Robots and Systems*, pp. 12 064–12 069.

[28] O. Scheel, L. Bergamini, M. Wolczyk, B. Osiński, and P. Ondruska, "Urban driver: Learning to drive from real-world demonstrations using policy gradients," in *Conference on Robot Learning*. PMLR, 2022, pp. 718–728.

[29] K. Renz, K. Chitta, O.-B. Mercea, A. Koepke, Z. Akata, and A. Geiger, "Plant: Explainable planning transformers via object-level representations," *arXiv preprint arXiv:2210.14222*, 2022.

[30] Y. Zheng, R. Liang, K. ZHENG, J. Zheng, L. Mao, J. Li, W. Gu, R. Ai, S. E. Li, X. Zhan *et al.*, "Diffusion-based planning for autonomous driving with flexible guidance," in *ICLR 2025 Workshop on Deep Generative Model in Machine Learning: Theory, Principle and Efficacy*.

[31] J. Cheng, Y. Chen, and Q. Chen, "Pluto: Pushing the limit of imitation learning-based planning for autonomous driving," *arXiv preprint arXiv:2404.14327*, 2024.

[32] B. Jaeger, D. Dauner, J. Beißwenger, S. Gerstenecker, K. Chitta, and A. Geiger, "Carl: Learning scalable planning policies with simple rewards," *arXiv preprint arXiv:2504.17838*, 2025.

[33] R. Chekroun, T. Gilles, M. Toromanoff, S. Hornauer, and F. Moutarde, "Mbappe: Mcts-built-around prediction for planning explicitly," in *IEEE IV Symposium 2024*, pp. 2062–2069.

[34] Z. Huang, P. Karkus, B. Ivanovic, Y. Chen, M. Pavone, and C. Lv, "Dtpp: Differentiable joint conditional prediction and cost evaluation for tree policy planning in autonomous driving," in *2024 IEEE International Conference on Robotics and Automation*, pp. 6806–6812.

[35] D. Dauner, M. Hallgarten, A. Geiger, and K. Chitta, "Supplementary material for parting with misconceptions about learning-based vehicle motion planning."

[36] Z. Huang, H. Liu, and C. Lv, "Gameformer: Game-theoretic modeling and learning of transformer-based interactive prediction and planning for autonomous driving," in *Proceedings of the IEEE International Conference on Computer Vision*, 2023, pp. 3903–3913.

[37] M. Hallgarten, M. Stoll, and A. Zell, "From prediction to planning with goal conditioned lane graph traversals," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems*, pp. 951–958.

[38] Q. Sun, S. Zhang, D. Ma, J. Shi, D. Li, S. Luo, Y. Wang, N. Xu, G. Cao, and H. Zhao, "Large trajectory models are scalable motion predictors and planners," *arXiv preprint arXiv:2310.19620*, 2023.

[39] J. Xu, B. Guo, X. Liu, W. Hong, L. Li, C. Xi, Y. Shi, P. Wang, and R. Di, "UniTSG: Unified modeling token scenarios generating for autonomous driving task," Waymo, Technical Report, 2025. [Online]. Available: https://storage.googleapis.com/waymo-uploads/files/research/2025%20Technical%20Reports/2025%20WOD%20Scenario%20Generation%20Challenge%20-%202nd%20Place%20-%20UniTSG.pdf