

ReflexDiffusion: Reflection-Enhanced Trajectory Planning for High-lateral-acceleration Scenarios in Autonomous Driving

Xuemei Yao¹, Xiao Yang^{2,3*}, Jianbin Sun^{1*}, Liuwei Xie^{2,3}, Xuebin Shao⁴,
Xiyu Fang⁴, Hang Su^{2,3}, Kewei Yang¹

¹College of Systems Engineering, National University of Defense Technology

²Department of Computer Science and Technology, Institute for AI, BNRist Center, Tsinghua University

³Tsinghua-Bosch Joint ML Center

⁴CATARC Intelligent Technology

{yaoxuemei2019,sunjianbin,kayyang27}@nudt.edu.cn, yangxiao19@tsinghua.org.cn, lxiear@outlook.com,
{shaoxuebin,fangxiyu}@catarc.ac.cn, suhangss@tsinghua.edu.cn

Abstract

Generating safe and reliable trajectories for autonomous vehicles in long-tail scenarios remains a significant challenge, particularly for **high-lateral-acceleration maneuvers** such as sharp turns, which represent critical safety situations. Existing trajectory planners exhibit systematic failures in these scenarios due to data imbalance. This results in insufficient modelling of vehicle dynamics, road geometry, and environmental constraints in high-risk situations, leading to suboptimal or unsafe trajectory prediction when vehicles operate near their physical limits. In this paper, we introduce **ReflexDiffusion**, a novel inference-stage framework that enhances diffusion-based trajectory planners through reflective adjustment. Our method introduces a gradient-based adjustment mechanism during the iterative denoising process: after each standard trajectory update, we compute the gradient between the conditional and unconditional noise predictions to explicitly amplify critical conditioning signals, including road curvature and lateral vehicle dynamics. This amplification enforces strict adherence to physical constraints, particularly improving stability during high-lateral-acceleration maneuvers where precise vehicle-road interaction is paramount. Evaluated on the nuPlan Test14-hard benchmark, ReflexDiffusion achieves a 14.1% improvement in driving score for high-lateral-acceleration scenarios over the state-of-the-art (SOTA) methods. This demonstrates that inference-time trajectory optimization can effectively compensate for training data sparsity by dynamically reinforcing safety-critical constraints near handling limits. The framework’s architecture-agnostic design enables direct deployment to existing diffusion-based planners, offering a practical solution for improving autonomous vehicle safety in challenging driving conditions.

Code —

<https://github.com/Luminous2028/ReflexDiffusion.git>

1 Introduction

Safe autonomous driving fundamentally depends on generating reliable trajectories in long-tail scenarios. High-lateral-acceleration maneuvers presenting a critical paradox: they

exhibit the highest accident risk yet receive the lowest representation in training data(Choudhari and Maji 2021; Antonios, Konstantinos, and Basil 2023; Rafiei, Abdi Kordani, and Zarei 2024; Kota et al. 2025; Sun et al. 2025). While recent advances have focused on generating synthetic edge cases to augment training datasets(Tang et al. 2025; Xu et al. 2025; Peng et al. 2025), a crucial gap remains in fully leveraging existing data to improve planning quality during inference. Current trajectory planning approaches fall into two categories: rule-based methods employ manually designed constraints(Dauner et al. 2023; Treiber, Hennecke, and Helbing 2000) but fail to adapt to novel scenarios; learning-based approaches that utilize imitation(Cheng, Chen, and Chen 2024; Cheng et al. 2024) or reinforcement learning(Liu, Zhang, and Zhao 2022; Liu et al. 2023) struggle to capture the multi-modal nature of real-world driving, often resulting in suboptimal planning policies.

Diffusion models have emerged as a promising solution for trajectory planning. For instance, the Diffusion Planner (Zheng et al. 2025) jointly models agent dynamics, lane topology, and navigation information using a transformer architecture. However, these approaches exhibit critical limitations in **high-lateral-acceleration scenarios**, where they often generate unsafe trajectories due to curvature-speed decoupling. Furthermore, they typically rely on carefully designed classifier guidance functions. These functions must be continuously differentiable while satisfying multiple constraints-such as passenger comfort and collision avoidance, significantly complicating deployment.

The field of large language models offers a compelling paradigm through iterative reflection mechanisms. In this approach, generate-evaluate-refine cycles dramatically enhance output quality for high-stakes tasks that require complex constraint satisfaction (Ji et al. 2023; Shinn et al. 2023; Madaan et al. 2023; Gou et al. 2024). This paradigm holds significant promise for high-lateral-acceleration trajectory planning, where vehicles must operate near their physical limits while maintaining strict adherence to centripetal force constraints. Although reflection-based approaches have shown success across numerous domains (Liu, Sferrazza, and Abbeel 2023; Bai et al. 2025), their application to trajectory planning remains unexplored. This

*Corresponding authors

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

gap presents an untapped opportunity to enhance safety when vehicles operate near their handling limits.

In this paper, we present ReflexDiffusion, a physics-aware reflection framework that pioneers a new approach for trajectory generation in autonomous driving. ReflexDiffusion operates through a dual-phase inference process: standard denoising generates an initial trajectory, which is then refined by a novel reflection phase that injects physics-aware gradients to explicitly enforce curvature-speed-acceleration coupling. This is achieved through conditional gradient ascent that amplifies critical physical constraints, iteratively refining the trajectory until convergence. Specifically, at each denoising step, we compute the gradient difference between conditional and unconditional noise predictions. This operation reveals how physical constraints influence trajectory generation. By strategically amplifying and injecting these gradients, we compel the trajectory to strictly adhere to centripetal force limitations, preventing dangerous speed-curvature violations that plague existing methods. Crucially, this approach eliminates the need for hand-crafted classifier guidance while maintaining full compatibility with existing diffusion-based planners, requiring no structural modifications. Thus, ReflexDiffusion fundamentally transforms the planning paradigm: rather than treating physical constraints as external penalties, it embeds them into the generative process through gradient-based reflection, enabling robust safety-critical planning tasks.

We demonstrate that ReflexDiffusion achieves state-of-the-art (SOTA) performance on high-lateral-acceleration scenarios in the nuPlan Test14-hard (Cheng et al. 2024) and Test14-random (Cheng et al. 2024) benchmarks. The method excels in safety-critical edge cases such as U-turns, where curvature-speed coupling is paramount. Our main contributions are as follows:

- **Pioneering physics-aware reflection for trajectory generation.** We introduce the first adaptation of the reflection mechanism to autonomous driving trajectory planning. Our framework leverages gradient-based adjustment to perform iterative trajectory correction during diffusion sampling, thereby addressing critical safety gaps in real-time planning.
- **Breakthrough performance in high-lateral-acceleration scenarios.** By explicitly amplifying curvature-speed coupling through physics-aware reflection, ReflexDiffusion achieves a **14.1% improvement** in driving score on high-lateral-acceleration scenarios within the Test14-hard benchmark, setting a new state-of-the-art.
- **Architecture-agnostic generalization.** As a plug-and-play module, the framework boosts the driving score of the Diffusion-es planner by **22.5%** on Test14-hard when deployed during inference, demonstrating compelling universality and transferability without any architectural modifications.

2 Related Works

Autonomous vehicles trajectory generation. Trajectory generation for autonomous vehicles is primarily addressed

by rule-based and learning-based methods. Rule-based approaches use explicit logical rules derived from traffic and safety knowledge (Treiber, Hennecke, and Helbing 2000; Dauner et al. 2023), but lack generalization in complex scenarios. Learning-based methods, such as imitation learning (Cheng, Chen, and Chen 2024) and reinforcement learning (Shalev-Shwartz, Shammah, and Shashua 2016; Liu et al. 2023), learn policies from data but often produce suboptimal, unimodal trajectories. Diffusion models address this by capturing multi-modal behavior, however, they frequently depend on post-processing or guidance functions (Yang et al. 2024; Zheng et al. 2025), which compromises their flexibility and interpretability. Our framework overcomes these limitations by eliminating the need for explicit guidance or rule-based refinement, instead enabling direct and flexible trajectory adaptation through a novel diffusion reflection mechanism.

Challenges in long-tail scenarios. Long-tail scenarios involving high lateral acceleration (e.g., emergency curve negotiation) pose significant challenges to existing planners. These challenges stem from data scarcity, the oversimplified constraints of rule-based methods, and the static risk adaptation in imitation learning (Becker et al. 2023; Yoshioka and Suzuki 2025). Existing approaches often fail to incorporate key physical principles, such as curvature-speed-acceleration coupling ($a_y = \kappa v^2$), resulting in unstable vehicle behavior near handling limits. Our method bridges this gap by leveraging inference-time reflection to dynamically reinforce κ - v relationship and adjust safety margins using real-time lateral acceleration feedback.

Reflection mechanisms. Reflection mechanisms iteratively refine outputs through generate-evaluate-adjust cycles, improving correctness and safety in domains such as large language models and computer vision (Liu, Sferazza, and Abbeel 2023; Madaan et al. 2023; Bai et al. 2025). However, they remain underexplored in trajectory planning, especially for high-dynamic scenarios. Existing adaptations in this domain are either computationally prohibitive for real-time use or fail to guarantee adherence to physical constraints (Yang et al. 2024; Zheng et al. 2025). Our work introduces a real-time reflection mechanism that employs physics-compliant gradient adjustment and safety-gated triggers to proactively correct trajectories upon detecting curvature-speed misalignment.

Generalization of trajectory generation methods. Generalization under distribution shift remains a persistent challenge. Rule-based methods typically require manual re-engineering per scenario, while data-driven generative models, such as GANs (Jiang et al. 2023) and the Diffusion Planner (Zheng et al. 2025), often lack inference-time adaptability. In contrast, our reflection framework is architecture-agnostic. This design enables zero-shot transfer to unseen planners and delivers improved performance on nuPlan, without any structural modifications to the base model.

3 Problem Formulation

Task Definition. High-lateral-acceleration scenarios represent a critical subset of long-tail driving situations characterized by extreme vehicle dynamics near handling limits.

Formally, a driving scenario \mathcal{S} is classified as high-lateral-acceleration (Karnchanachari et al. 2024) if:

$$\exists t \in [t_0, t_0 + \Delta t], \quad |a_y(t)| \geq a_{th}, \quad \Delta t \geq t_{min} \quad (1)$$

where $a_y(t)$ is the instantaneous lateral acceleration (m/s²); $a_{th} = 4.0$ m/s² is the acceleration threshold; $\Delta t \geq 0.5$ s is the minimum duration (nuPlan benchmark).

Conventional planners exhibit critical failures in these scenarios due to curvature-speed decoupling. The curvature κ of the planned trajectory and the vehicle speed v fail to satisfy the centripetal force constraint $a_y = \kappa v^2$. A violation occurs when $|\kappa v^2 - a_y| > 4.0$ m/s², $\kappa = 1/R$, resulting in understeer or oversteer, where κ represents planned trajectory curvature and R represents turning radius.

Our objective is to mitigate this decoupling problem by explicitly reinforcing the physical coupling between curvature, velocity, and acceleration through the reflection mechanism, using the same joint modelling approach as in Diffusion Planner (Zheng et al. 2025), which enables interaction-aware planning.

Inputs. The inputs comprise (1) Ego state: Current position (x, y) , heading $(\cos \theta, \sin \theta)$; (2) Agent history: Neighbor agents' states over $L = 21$ past timesteps, and their attributes include type (car, bicycle and pedestrian), size and velocity. (3) Scene context; (4) Vectorized HD Map: Lane polylines with traffic lights and speed limits; (5) Static obstacles: Construction zones, road barriers; (6) Route navigation: Sequence of route lanes.

Outputs. The output formation is a joint trajectory matrix including ego trajectory as well as M neighbors trajectories in the next 8 seconds at 10Hz, i.e., $\mathbf{x}^{(0)} \in \mathbb{R}^{(M+1) \times \tau \times 4}$:

$$\mathbf{x}^{(0)} = \begin{bmatrix} x_{ego}^1 & \cdots & x_{ego}^\tau \\ x_{neigh_1}^1 & \cdots & x_{neigh_1}^\tau \\ \vdots & \ddots & \vdots \\ x_{neigh_M}^1 & \cdots & x_{neigh_M}^\tau \end{bmatrix}, \quad x_{agent}^k = \begin{bmatrix} x \\ y \\ \cos \theta \\ \sin \theta \end{bmatrix}$$

4 Methodology

In this section, we present ReflexDiffusion, which contains four modules to capture and make the most of the underlying characteristics of the long-tail data in training sets, as illustrated in Figure 1.

4.1 Training Module

To address data scarcity in long-tail scenarios, we introduce a conditional dropout strategy (Ho and Salimans 2022) into the Diffusion Planner (Zheng et al. 2025) training framework. While preserving its unified modeling capacity for neighbor vehicles, lane structures, and navigation signals, our approach strategically drops critical physics-coupled features during training to simulate real-world sensory degradation. Crucially, this includes randomly masking road turning radius R and vehicle speed v to simulate scenarios where their coupling $a_y = v^2/R$ is obscured. This forces the model to learn robust representations under partial observability, directly addressing the physics-disalignment challenge in extreme maneuvers. Given a

training sample with (1) \mathbf{x}_{gt} : Ground-truth trajectory; (2) $\mathbf{c}_{full} = [\mathbf{c}_{neighbors}, \mathbf{c}_{lanes}, \mathbf{c}_{nav}, \mathbf{c}_{static_obj}]$: Condition vector (neighbor vehicle data, lanes, navigation information, static objects); (3) $\mathbf{c}_{decouple} = [\mathbf{c}_{nav}]$, which drops lane information such as road turning radius R and agent data including ego vehicle speed v .

We generate masked condition $\tilde{\mathbf{c}}$ via 10% conditional dropout:

$$\tilde{\mathbf{c}} = \begin{cases} \mathbf{c}_{decouple} & \text{with probability } p_{drop} \\ \mathbf{c}_{full} & \text{otherwise} \end{cases} \quad (2)$$

The diffusion training objective then becomes:

$$\mathcal{L}_\theta = \mathbb{E}_{t \sim \mathcal{U}(0, T), \mathbf{x}_t \sim q_{t_0}(\mathbf{x}_t | \mathbf{x}_{gt}, \tilde{\mathbf{c}})} [\|\mathbf{x}_{gt} - \mathbf{x}_t\|^2] \quad (3)$$

while the diffusion training objective in Diffusion Planner is:

$$\mathcal{L}_\theta = \mathbb{E}_{t \sim \mathcal{U}(0, T), \mathbf{x}_t \sim q_{t_0}(\mathbf{x}_t | \mathbf{x}_{gt}, \mathbf{c}_{full})} \|\mathbf{x}_{gt} - \mathbf{x}_t\|^2 \quad (4)$$

4.2 Denoising Module

To address trajectory quality degradation in long-tail scenarios, particularly the inability to handle non-differentiable safety constraints, existing SOTA Diffusion Planners (Zheng et al. 2025) rely on complex classifier guidance (CG) during denoising. This approach requires handcrafting a guidance function $f_{cg}(\tau)$ that must be continuously differentiable, multi-objective and computationally tractable.

However, these requirements lead to oversimplified approximations, severely limiting performance in safety-critical edge cases. Therefore, we utilize classifier-free guidance (CFG) (Ho and Salimans 2022) as a principled alternative. The key innovation lies in embedding trajectory curvature constraint awareness directly into the diffusion model through conditional dropout during training, eliminating the need for auxiliary guidance functions. The CFG denoising process is formalized as:

$$\epsilon_\theta^t(\mathbf{x}_t) = \epsilon_\theta(\mathbf{x}_t | \mathbf{c}_{decouple}) + \lambda_1 [\epsilon_\theta(\mathbf{x}_t | \mathbf{c}_{full}) - \epsilon_\theta(\mathbf{x}_t | \mathbf{c}_{decouple})] \quad (5)$$

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\mathbf{x}_t - \frac{\beta_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_\theta^t(\mathbf{x}_t) + \sigma_t \mathbf{z} \right) \quad (6)$$

where $\epsilon_\theta^t(\mathbf{x}_t)$ represents the noise predicted by the CFG at time t , λ_1 is the guidance scale controlling constraint adherence intensity, $\beta_t = 1 - \alpha_t$, $\bar{\alpha}_t = \alpha_1 \alpha_2 \dots \alpha_t$, and $\mathbf{z} \sim \mathcal{N}(0, 1)$.

4.3 Reflection Module

Inspired by the reflective reasoning in large language models (LLMs) that iteratively refine outputs via self-critique (Ji et al. 2023; Madaan et al. 2023), we are the first to propose the reflection mechanism for autonomous driving trajectory generation. This paradigm shift addresses inadequate long-tail information utilization in standard diffusion denoising, where single-pass sampling fails to recover critical scene semantics in rare events. Conventional denoising suffers from (1) Information dilution: Progressive noise removal attenuates long-tail features. (2) Error accumulation: Early-stage

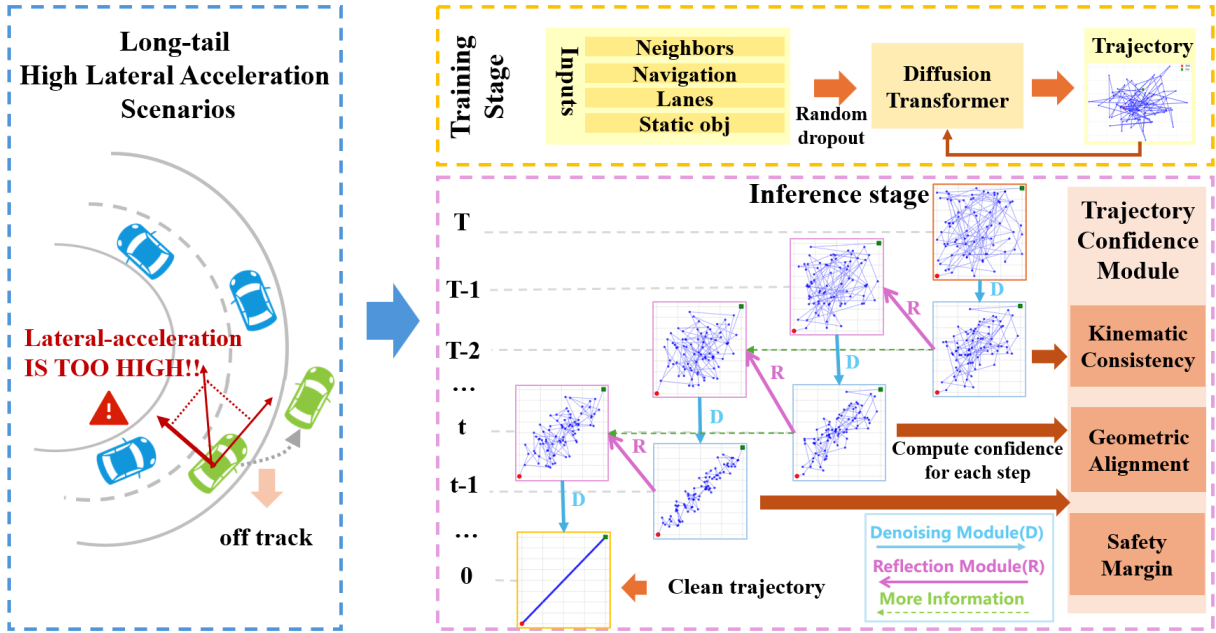


Figure 1: Architecture of ReflexDiffusion. (a) Training Module enhances model robustness to physically-coupled features via a conditional dropout strategy. (b) Denoising Module employs classifier-free guidance to generate initial trajectories. (c) Reflection Module iteratively refines trajectories by injecting physics-aware gradients to enforce curvature-speed-acceleration coupling constraints during inference. (d) Trajectory Confidence Module assesses trajectory reliability, and dynamically triggers the reflection mechanism to ensure safety.

deviations amplify in later steps. (3) No recourse: Once generated, low-quality trajectories cannot be revised.

Our physics-aware reflection mechanism breaks this linear flow through iterative information boosting. After obtaining x_{t-1} from standard CFG denoising (Eq. 6), to make the next reflection deterministic, we use the DDIM scheduler (Song, Meng, and Ermon 2021) for denoising, for which the denoising formula becomes:

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\mathbf{x}_t - \frac{\beta_t}{\sqrt{\alpha_t - \bar{\alpha}_t + \sqrt{1 - \bar{\alpha}_t}}} \epsilon_{\theta}^t(\mathbf{x}_t) \right) \quad (7)$$

After obtaining \mathbf{x}_{t-1} from CFG denoising, we compute trajectory confidence $\mathcal{C}(\mathbf{x}_{t-1})$. If $\mathcal{C}(\mathbf{x}_{t-1}) < \gamma$, reflection mechanism is triggered. We approximate the noise predicted at timestep t with timestep $t-1$ along the reflection path, i.e, set $\epsilon_{\theta}^t(\mathbf{x}_t) \approx \epsilon_{\theta}^t(\mathbf{x}_{t-1})$.

$$\mathbf{x}'_t = \sqrt{\alpha_t} \mathbf{x}_{t-1} + \frac{\beta_t}{\sqrt{\alpha_t - \bar{\alpha}_t + \sqrt{1 - \bar{\alpha}_t}}} \epsilon_{\theta}^t(\mathbf{x}_{t-1}) \quad (8)$$

$\epsilon_{\theta}^t(\mathbf{x}_{t-1})$ contains information about the coupling of road curvature κ and vehicle speeds v , which can be rewritten as Δ_{couple} :

$$\begin{aligned} \Delta_{couple} = & \underbrace{\epsilon_{\theta}(\mathbf{x}_{t-1} | \mathbf{c}_{decouple})}_{\text{exploration}} \\ & + \lambda_2 \underbrace{(\epsilon_{\theta}(\mathbf{x}_{t-1} | \mathbf{c}_{full}) - \epsilon_{\theta}(\mathbf{x}_{t-1} | \mathbf{c}_{decouple}))}_{\text{condition gradient}} \end{aligned} \quad (9)$$

where λ_2 is the reflection guidance scale controlling constraint adherence intensity.

Projecting Δ_{couple} onto centripetal constrained manifolds:

$$\Delta_{proj} = \mathbf{P} \cdot \Delta_{couple}, \quad \mathbf{P} = \begin{bmatrix} \frac{\partial(\kappa v^2)}{\partial \kappa} & \frac{\partial(\kappa v^2)}{\partial v} \\ 0 & 1 \end{bmatrix} \quad (10)$$

The projection matrix \mathbf{P} enforces physical constraints by mapping gradient adjustments Δ_{couple} onto the centripetal force manifold $a_y \approx \kappa v^2$. The projection matrix \mathbf{P} has the structure:

- Top row ($v^2, 2\kappa v$): Amplifies curvature (κ)-velocity (v) coupling using partial derivatives of κv^2 .
- Bottom row ($0, 1$): Preserves unconstrained motion freedom. This ensures trajectory corrections strictly satisfy vehicle dynamics while maintaining distributional coherence through iterative denoising:

$$\mathbf{x}'_t = \sqrt{\alpha_t} \mathbf{x}_{t-1} + b \cdot \Delta_{proj} \quad (11)$$

The reflection sample \mathbf{x}'_t adds more information about road curvature κ and vehicle speeds v compared to the pre-reflection sample \mathbf{x}_t :

$$\mathbf{x}'_t = \mathbf{x}_t - \frac{\beta_t}{\sqrt{\alpha_t - \bar{\alpha}_t + \sqrt{1 - \bar{\alpha}_t}}} \epsilon_{\theta}^t(\mathbf{x}_t) + b \cdot \Delta_{proj} \quad (12)$$

Compared to prior works, our reflection mechanism provides distinct advantages. Unlike classifier guidance (CG) or classifier-free guidance (CFG) methods that rely on one-shot generation, limiting their capacity for error correction,

Type	Planner	Test14-Hard		Test14-Random	
		NR	R	NR	R
Rule-based	IDM	36.74	62.42	67.61	64.66
	PDM-Closed	32.55	53.03	75.69	82.17
Hybrid	PDM-Hybrid	32.54	53.04	75.97	82.17
	Gameformer	53.12	57.46	82.60	79.49
	SAH-Drive	43.08	57.40	91.18	89.27
Learning-based	UrbanDriver	26.09	14.66	29.17	33.27
	Diffusion-es	44.63	52.72	88.20	84.20
	PlanCNN	20.16	28.03	48.20	51.38
	PlanTF	38.24	42.52	73.66	67.81
	Pluto	42.21	45.98	81.67	75.95
	Diffusion Planner	58.47	57.41	71.60	82.88
	Diffusion Planner w/ conditional dropout	12.60	14.04	38.45	44.88
	Diffusion Planner w/ conditional dropout+cfg	44.40	55.55	76.44	70.21
	Ours	59.94	65.53	86.40	71.57

Table 1: Closed-loop planning results on high-lateral-acceleration scenarios in the long-tail dataset and regular dataset. To ensure that the performance gains are solely attributed to our proposed reflection mechanism and not training strategies, we retrain the current SOTA baseline Diffusion Planner, under two additional settings: Conditional dropout and Classifier-free guidance (CFG) for a comprehensive and fair comparison. Our method achieves SOTA on Test14-hard high-lateral-acceleration scenarios. NR: non-reactive mode. R: reactive mode.

our approach enables iterative refinement through cyclic noise reversion. Whereas RL-based planners require costly retraining to adapt to new scenarios, our solution operates purely during inference without parameter updates. While ensemble methods achieve robustness via multiple models at the expense of high computational overhead, our method maintains efficiency by utilizing a single model architecture with minimal latency increase.

4.4 Trajectory Confidence Module

To assess the reliability of generated trajectories throughout the denoising and reflection cycles, we introduce a multi-factor confidence metric. This module fulfills two key roles: (1) determining when to trigger reflection, and (2) ensuring safety through real-time risk evaluation to enable effective fallback mechanisms.

The confidence metric $\mathcal{C}(\mathbf{x}_t)$ quantifies the reliability of generated trajectories in high-lateral-acceleration scenarios, integrating three domain-specific factors:

1. Kinematic Consistency (\mathcal{D}_{kin}). Assesses compliance with fundamental motion constraints:

$$\mathcal{D}_{\text{kin}} = \exp(-m_1 |a_y^{\text{traj}} - a_y^{\text{ref}}|) \cdot \sigma(j_{\text{max}} - |j_{\text{lat}}|) \quad (13)$$

$$a_y^{\text{traj}} = \kappa_\tau v^2, j_{\text{lat}} = \frac{da_y}{dt}$$

2. Geometric Alignment ($\mathcal{G}_{\text{align}}$). Quantifies adherence to road geometry:

$$\mathcal{G}_{\text{align}} = \mathcal{R}_{\text{curv}} \times \mathcal{R}_{\text{dev}}$$

$$\mathcal{R}_{\text{curv}} = \exp(-m_2 |\kappa_\tau - \kappa_{\text{road}}| \cdot R_{\text{curve}}) \quad (14)$$

$$\mathcal{R}_{\text{dev}} = 1 - \frac{\max(0, d_{\text{dev}} - d_{\text{safe}})}{d_{\text{max}}}$$

where $R_{\text{curve}} = 1/|\kappa_{\text{road}}|$: Current road curve radius, $d_{\text{dev}} = \max \|\mathbf{p}_\tau - \mathbf{p}_{\text{center}}\|_2$: Maximum lateral deviation.

3. Safety Margin ($\mathcal{S}_{\text{margin}}$). Evaluates risk buffers:

$$\mathcal{S}_{\text{margin}} = \sigma\left(\frac{\text{TTC} - 2.5}{0.5}\right) \cdot (1 - p_{\text{ODA}}) \cdot \cos(\Delta\psi) \quad (15)$$

where TTC represents time-to-collision with nearest obstacle, p_{ODA} is probability of out-of-drivable-area violation, and $\Delta\psi$ is heading deviation from road tangent.

5 Experiments

In this section, we conduct extensive experiments to demonstrate the performance of our method.

5.1 Experimental Setup

Datasets. To assess the performance of ReflexDiffusion on high-lateral-acceleration scenarios in long-tail dataset, we utilize Test14-hard (Cheng et al. 2024) and Test14-random (Cheng et al. 2024) as our benchmarks. Both benchmarks are developed to cover the 14 scenario types including our target scenarios in the nuPlan Challenge, with the former introduced especially to represent long-tail scenarios.

Baselines. The baselines are mainly divided into three groups (Dauner et al. 2023): Rule-based (Treiber, Hennecke, and Helbing 2000; Dauner et al. 2023), Learning-based (Scheel, Bergamini, and Wołczyk 2021; Yang et al. 2024; Renz et al. 2022; Cheng et al. 2024; Cheng, Chen, and Chen 2024; Zheng et al. 2025), and Hybrid methods (Dauner et al. 2023; Huang, Liu, and Lv 2023; Fan et al. 2025), which incorporate additional refinement to the outputs of the learning-based model. We compare ReflexDiffusion against the baselines above, which are detailed in Appendix.

5.2 Main Results

We compare the performance of ReflexDiffusion with baselines on high-lateral-acceleration scenarios in the long-tail

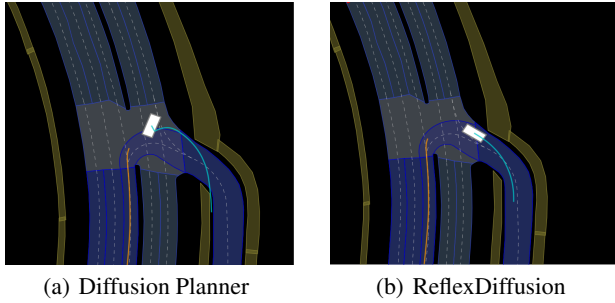


Figure 2: Visualization comparison in U-turn scenario from high-lateral-acceleration scenarios. The Diffusion Planner’s trajectory veers out of the lane during the turn, while ReflexDiffusion makes the turn without incident!

Planner	per-step	e2e
Diffusion-es	/	7612.7
Diffusion Planner	/	35.7
ReflexDiffusion w/o reflection	3.3	35.7
Ours	6.3	36.1

Table 2: Runtime tests among diffusion-based planners. The runtime of ReflexDiffusion stays near the current SOTA baseline, and trades minor latency for higher robustness.

dataset Test14-hard, as well as the regular dataset Test14-random. The quantitative and qualitative results are shown in Table 1 to Table 6 and Figure 2 to Figure 3, from which we derive the following findings.

Our method achieves state-of-the-art (SOTA) performance on high-lateral-acceleration scenarios in long-tail benchmarks while maintaining strong performance on regular datasets. As summarized in Table 1, ReflexDiffusion outperforms the current SOTA Diffusion Planner on the Test14-hard benchmark in both non-reactive and reactive modes, achieving a significant 14.1% improvement in driving score for the reactive mode. Concurrently, performance on the regular dataset is also substantially enhanced, with a 20.7% improvement in the non-reactive mode. Runtime analysis conducted on a single RTX 4090 GPU (Table 2) indicates that enabling the reflection mechanism increases the per-step latency from 3.3 ms to 6.3 ms. However, since reflection is triggered in only $\leq 0.5\%$ of evaluated real-world driving cases, the average runtime is calculated as $0.5\% \times 122.7 + 99.5\% \times 35.7 \approx 36.1$ ms. This remains comparable to the SOTA baseline, sustaining a control frequency of > 20 Hz suitable for real-time deployment. Crucially, this marginal increase in latency is traded for a substantial gain in planning robustness and safety.

The effectiveness of our method is further demonstrated through trajectory confidence. The superior performance of ReflexDiffusion in high-lateral-acceleration scenarios is corroborated by trajectory confidence analysis. In a representative U-turn scenario where the baseline Diffusion Planner fails, our method achieves a significantly higher driving score (Table 3). A qualitative comparison in Figure 2

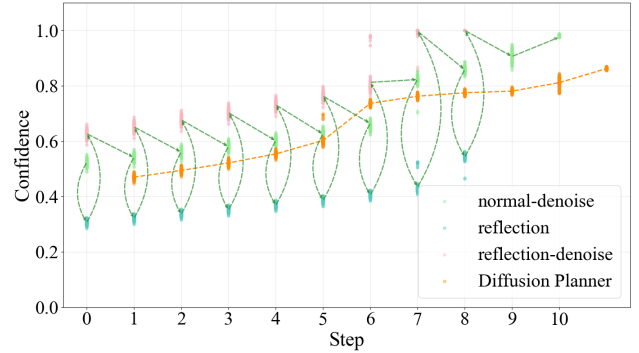


Figure 3: Visualization of trajectory confidence in inference process. --- represents the trend of trajectory confidence for ReflexDiffusion and --- represents the trend of trajectory confidence for Diffusion Planner.

Planner	Score	Trajectory Confidence		
		nor	re	re-de
Diffusion Planner	0.00	0.48	/	/
Ours	100.00	0.87	0.36	0.73

Table 3: Driving score comparisons on U-turn scenario from high-lateral-acceleration scenarios. Including average trajectory confidence of normal-denoising (nor), reflection (re) and reflection-denoising (re-de) process.

illustrates this improvement: while the baseline planner generates trajectories that violate drivable areas due to excessive lateral acceleration, ReflexDiffusion produces trajectories that maintain both safety and feasibility.

We attribute the baseline’s failures to an insufficient representation of extreme vehicle dynamics in the training data, which critically impairs generalization to high-lateral-acceleration cases. To address this, ReflexDiffusion introduces a reflective adjustment step during inference. After each denoising iteration, a physics-aware gradient—derived from the difference between conditional and unconditional predictions—is computed and injected to explicitly amplify critical physical constraints. Through this iterative refinement, trajectories progressively converge to satisfy the centripetal force constraint $a_y \leq v^2/R$. As visualized in the confidence plot (Figure 3), the trajectory confidence temporarily decreases during the reflection phase due to exploratory gradient adjustments. This dip signifies an active optimization process against physical constraints. The confidence subsequently recovers and rises to a peak level after the final denoising step. This pattern demonstrates how the reflection mechanism compensates for training data scarcity by dynamically reinforcing vehicle-road interaction physics near handling limits, effectively converting transient uncertainty into enhanced safety margins.

Our framework demonstrates strong generalizability across different diffusion-based planners. Our framework is architecture-agnostic, requiring no structural changes to the base planner. Its core innovation lies in an inference-stage reflection mechanism, with the only training modifi-

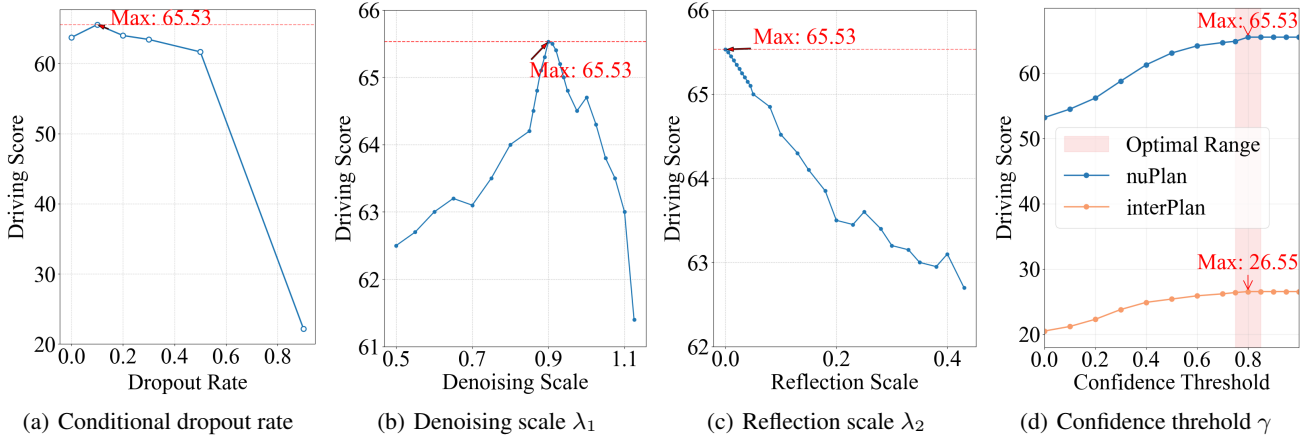


Figure 4: Ablation Studies. The optimal values for the four parameters: conditional dropout rate, denoising scale λ_1 , reflection scale λ_2 and confidence threshold γ are 0.1, 0.9, 0.0, and 0.8, respectively.

Planner	Test14-hard driving score	Test14-random driving score
Diffusion Planner	57.41	71.60
w/ ReflexDiffusion	65.53	86.40

Table 4: Generalizability test on Diffusion Planner. The result proves that ReflexDiffusion can significantly enhance the performance of Diffusion Planner in high-risk scenarios.

Planner	Test14-hard driving score	Test14-random driving score
Diffusion-es	31.88	/
w/ ReflexDiffusion	39.04	/

Table 5: Generalizability test on Diffusion-es. The result indicates that the enhancement effect of ReflexDiffusion is universal and it is a general reasoning stage optimization scheme independent of the model architecture.

cation being a random dropout of input features. This design enables seamless transfer across different diffusion-based planners. Generalization tests on both Diffusion Planner (Zheng et al. 2025) and Diffusion-es (Yang et al. 2024) confirm its effectiveness, demonstrating significant performance improvements in long-tail scenarios as shown in Table 4 and Table 5. We train the Diffusion-es conditional diffusion model for **50 epochs** as a baseline, and use the same settings to train a new model with a conditional dropout rate of 10%. These results validate ReflexDiffusion as a general and practical strategy for enhancing the safety and robustness of existing trajectory planners.

5.3 Ablation Studies

Design Choices for training. We demonstrate the effectiveness of conditional dropout in the training stage, as shown in Table 6, and its optimal dropout rate is demonstrated in Figure 4(a). The results show that the conditional dropout is

Type	Planner	Score
Base	ReflexDiffusion	65.53
	w/o conditional dropout	23.86
Simplification	w/o cfg denoising	59.85
	w/o reflection	53.21

Table 6: Ablation of each module during the training and inference process. The result clearly demonstrates that each module is crucial for ReflexDiffusion.

significant and the optimal performance is achieved at 10%. **Design Choices for inference.** We ablate the classifier-free guidance denoising and reflection mechanism, as shown in Table 6, as well as two hyperparameters: denoising scale λ_1 and reflection scale λ_2 shown in Figure 4(b) and 4(c), respectively. The results indicate that both classifier-free guidance and reflection mechanisms play a prominent role in enhancing the performance on long-tail datasets, with the optimal combination $(\lambda_1, \lambda_2)=(0.9, 0)$. Moreover, the confidence threshold γ is empirically calibrated via cross-dataset nuPlan and interPlan (Hallgarten et al. 2024) validation, as shown in Figure 4(d), giving a sensitivity range $[0.75, 0.85]$, balancing robustness and responsiveness.

6 Conclusion

This paper presents ReflexDiffusion, a novel framework that pioneers inference-stage reflection for diffusion-based trajectory planning. Inspired by the generate-evaluate-refine paradigm in large language models, we develop physics-aware reflection through conditional gradient injection during diffusion sampling. This innovation explicitly amplifies critical physical couplings, particularly the curvature-speed-acceleration relationship, enabling real-time trajectory correction without classifier guidance. Comprehensive nuPlan evaluations demonstrate significant advancements.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (NSFC) under grants Nos. 72231011, 72471238, 92370124, 92248303, 62276149.

References

- Antonios, T. E.; Konstantinos, A.; and Basil, P. 2023. Vehicles lateral acceleration and speed profiles investigation at the entry area of interchange ramps as a criterion of geometric road design. *Transportation research procedia*, 69: 13–20.
- Bai, L.; Shao, S.; Zhou, Z.; Qi, Z.; Xu, Z.; Xiong, H.; and Xie, Z. 2025. Zigzag Diffusion Sampling: Diffusion Models Can Self-Improve via Self-Reflection. In *ICLR*.
- Becker, J.; Imholz, N.; Schwarzenbach, L.; Ghignone, E.; Baumann, N.; and Magno, M. 2023. Model- and Acceleration-based Pursuit Controller for High-Performance Autonomous Racing. ArXiv:2209.04346.
- Cheng, J.; Chen, Y.; and Chen, Q. 2024. PLUTO: Pushing the Limit of Imitation Learning-based Planning for Autonomous Driving. ArXiv:2404.14327.
- Cheng, J.; Chen, Y.; Mei, X.; Yang, B.; Li, B.; and Liu, M. 2024. Rethinking Imitation-based Planners for Autonomous Driving. In *ICRA*.
- Choudhari, T.; and Maji, A. 2021. Risk Assessment of Horizontal Curves Based on Lateral Acceleration Index: A Driving Simulator-Based Study. *Transportation in Developing Economies*.
- Dauner, D.; Hallgarten, M.; Geiger, A.; and Chitta, K. 2023. Parting with Misconceptions about Learning-based Vehicle Motion Planning. In *CoRL*.
- Fan, Y.; Cui, Z.; Li, Z.; Ren, Y.; and Yu, H. 2025. SAH-Drive: A Scenario-Aware Hybrid Planner for Closed-Loop Vehicle Trajectory Generation. In *ICML*.
- Gou, Z.; Shao, Z.; Gong, Y.; Shen, Y.; Yang, Y.; Duan, N.; and Chen, W. 2024. CRITIC: Large Language Models Can Self-Correct with Tool-Interactive Critiquing. ArXiv:2305.11738.
- Hallgarten, M.; Zapata, J.; Stoll, M.; Renz, K.; and Zell, A. 2024. Can Vehicle Motion Planning Generalize to Realistic Long-tail Scenarios? In *IROS*.
- Ho, J.; and Salimans, T. 2022. Classifier-Free Diffusion Guidance. ArXiv:2207.12598.
- Huang, Z.; Liu, H.; and Lv, C. 2023. GameFormer: Game-theoretic Modeling and Learning of Transformer-based Interactive Prediction and Planning for Autonomous Driving. In *ICCV*.
- Ji, Z.; Yu, T.; Xu, Y.; Lee, N.; Ishii, E.; and Fung, P. 2023. Towards Mitigating Hallucination in Large Language Models via Self-Reflection. ArXiv:2310.06271.
- Jiang, W.; Zhao, W. X.; Wang, J.; and Jiang, J. 2023. Continuous Trajectory Generation Based on Two-Stage GAN. In *AAAI*.
- Karnchanachari, N.; Geromichalos, D.; Tan, K. S.; Li, N.; and Eriksen, C. 2024. Towards learning-based planning: The nuPlan benchmark for real-world autonomous driving. In *ICRA*.
- Kota, R. B.; Mehar, A.; Shankar, K. V. R. R.; and Pallela, S. S. 2025. Analyzing acceleration characteristics of right-turning vehicles at uncontrolled intersections: a study on longitudinal and lateral dynamics. *Innovative Infrastructure Solutions*.
- Liu, H.; Sferrazza, C.; and Abbeel, P. 2023. Chain of Hindsight Aligns Language Models with Feedback. ArXiv:2302.02676.
- Liu, J.; Hang, P.; qi, X.; Wang, J.; and Sun, J. 2023. MTD-GPT: A Multi-Task Decision-Making GPT Model for Autonomous Driving at Unsignalized Intersections. ArXiv:2307.16118.
- Liu, Y.; Zhang, Q.; and Zhao, D. 2022. Multi-task Safe Reinforcement Learning for Navigating Intersections in Dense Traffic. ArXiv:2202.09644.
- Madaan, A.; Tandon, N.; Gupta, P.; Hallinan, S.; Gao, L.; Wiegrefe, S.; Alon, U.; Dziri, N.; Prabhunoye, S.; Yang, Y.; Gupta, S.; Majumder, B. P.; Hermann, K.; Welleck, S.; Yazdanbakhsh, A.; and Clark, P. 2023. Self-Refine: Iterative Refinement with Self-Feedback. ArXiv:2303.17651.
- Peng, M.; Yao, R.; Guo, X.; Xie, Y.; Chen, X.; and Ma, J. 2025. Safety-Critical Traffic Simulation with Guided Latent Diffusion Model. ArXiv:2505.00515.
- Rafiei, A.; Abdi Kordani, A.; and Zarei, M. 2024. The Effect of Rutting Damage on Safety of Roadways in Horizontal Curves: With Focus on Lateral Acceleration. *International Journal of Pavement Research and Technology*.
- Renz, K.; Chitta, K.; Mercea, O.-B.; Koepke, A. S.; Akata, Z.; and Geiger, A. 2022. PlanT: Explainable Planning Transformers via Object-Level Representations. In *CoRL*.
- Scheel, O.; Bergamini, L.; and Wołczyk, M. 2021. Urban Driver: Learning to Drive from Real-world Demonstrations Using Policy Gradients. In *CoRL*.
- Shalev-Shwartz, S.; Shammah, S.; and Shashua, A. 2016. Safe, Multi-Agent, Reinforcement Learning for Autonomous Driving. ArXiv:1610.03295.
- Shinn, N.; Cassano, F.; Berman, E.; Gopinath, A.; Narasimhan, K.; and Yao, S. 2023. Reflexion: Language Agents with Verbal Reinforcement Learning. ArXiv:2303.11366.
- Song, J.; Meng, C.; and Ermon, S. 2021. Denoising Diffusion Implicit Models. In *ICLR*.
- Sun, F.; Li, K.; Chen, G.; Guo, D.; Kang, J.; and Ren, C. 2025. Development of combined simulation device for vehicle lateral acceleration and yaw. In *Proceedings of the Institution of Mechanical Engineers*.
- Tang, T.; Wei, D.; Jia, Z.; Gao, T.; Cai, C.; Hou, C.; Jia, P.; Zhan, K.; Sun, H.; JingChen, F.; Zhao, Y.; Liang, X.; Lang, X.; and Wang, Y. 2025. BEV-TSR: Text-Scene Retrieval in BEV Space for Autonomous Driving. In *AAAI*.
- Treiber, M.; Hennecke, A.; and Helbing, D. 2000. Congested traffic states in empirical observations and microscopic simulations. *Physical review E*.

Xu, C.; Petiushko, A.; Zhao, D.; and Li, B. 2025. Diff-Scene: Diffusion-Based Safety-Critical Scenario Generation for Autonomous Vehicles. In *AAAI*.

Yang, B.; Su, H.; Gkanatsios, N.; Ke, T.-W.; Jain, A.; Schneider, J.; and Fragkiadaki, K. 2024. Diffusion-ES: Gradient-Free Planning with Diffusion for Autonomous and Instruction-Guided Driving. In *CVPR*.

Yoshioka, T.; and Suzuki, K. 2025. Path and speed planning with longitudinal and lateral motion coupling on arbitrary road profile for driving assistance and automated driving systems. *Mechanical Engineering Journal*, 12(3): 24–00393.

Zheng, Y.; Liang, R.; Zheng, K.; Zheng, J.; Mao, L.; Li, J.; Gu, W.; Ai, R.; Li, S. E.; Zhan, X.; and Liu, J. 2025. Diffusion-Based Planning for Autonomous Driving with Flexible Guidance. In *ICLR*.