SCRP: REPURPOSING TRAJECTORY PREDICTION FOR SAFETY-CRITICAL RELEVANCE PLANNING WITH AGENT SELECTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Trajectory planning for autonomous driving requires reasoning about interactions among multiple agents under real-time constraints. Existing approaches often rely on increased model complexity and iterative trajectory generation to achieve strong performance, but incur high computational cost and limit scalability. In contrast, one-shot planners generate trajectories in a single forward pass and improve efficiency, yet fail to match iterative methods in complex scenarios. We identify that this performance gap arises from the absence of decision-aware interaction reasoning rather than limitations in model capacity or inference style. This paper presents a safety-critical relevance planning (SCRP) framework that enables efficient and high-performance one-shot planning. SCRP builds upon a generic *state encoder–interaction encoder–trajectory decoder* backbone inherited from trajectory prediction models and enables decision-aware interaction reasoning through two components: perceived safety modeling and risk-conditioned adaptive agent selection. A perceived safety modeling mechanism learns temporally evolving interaction risk and distills future-supervised safety cues into the latent interaction representation. In addition, a risk-conditioned adaptive selection mechanism allocates a scene-dependent interaction cardinality by retaining the most relevant agents under a risk-aware residual-mass criterion. The resulting interaction set adapts its capacity to scene complexity while suppressing weakly relevant context and preserving computational efficiency. Extensive experiments on large-scale autonomous driving benchmarks, including Waymo and nuPlan, demonstrate that SCRP consistently outperforms state-of-the-art (SOTA) planning models, particularly in challenging interactive and safety-critical scenarios. Adding SCRP reduces inference latency by 4.5%–9.1% relative to the corresponding prediction backbones, improving planning quality while retaining an efficient one-shot inference pipeline.

1 INTRODUCTION

Trajectory planning for autonomous driving requires the ego vehicle to generate safe and feasible motions under complex multi-agent interactions while satisfying strict real-time constraints Hu et al. (2023); Zheng et al. (2024); Zhang et al. (2025b); Huang et al. (2025). As a core component of the decision-making system, planning directly affects safety, traffic efficiency, and comfort. A central challenge is to reason about how surrounding agents may influence the ego vehicle in the near future and to produce a timely decision in dense interactive scenes.

Recent progress has substantially improved planning performance, but strong results often come with considerable computational or optimization overhead. Diffusion-based planners generate trajectories through iterative refinement Yang et al. (2024); Fan et al. (2025); Zheng et al. (2025). Although effective in complex scenarios, they require multiple forward passes at test time, which increases inference latency and limits real-time applicability. Reinforcement learning-based planners follow a different route by learning planning policies through task-specific policy optimization Zhang et al. (2025a); Liu et al. (2026); Li et al. (2026). While such methods can achieve strong performance, they typically depend on specialized training pipelines, careful reward design, and high optimization cost, which makes them heavy to train and adapt, as depicted in Figure 1(a). These observations motivate a

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

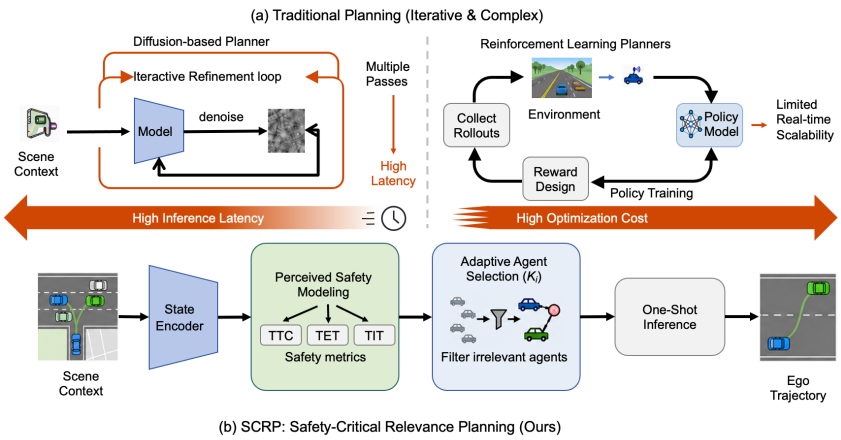


Figure 1: **Comparison between complex model-based planning and the proposed SCRP paradigm.** (a) Diffusion-based planners rely on iterative inference, while reinforcement learning-based planners typically require specialized policy optimization and high training cost; both introduce substantial overhead in practical deployment. (b) SCRP repurposes a trajectory prediction model for planning by introducing perceived safety modeling and risk-conditioned adaptive agent selection, enabling fast generation of safety-aware trajectories with minimal additional overhead.

more fundamental question: can we achieve strong interactive planning performance without relying on either iterative trajectory generation or planning-specific policy optimization?

A promising direction is to repurpose trajectory prediction models for planning. Modern trajectory prediction models already perform efficient one-shot reasoning over multi-agent scenes: they encode agent and map context, model interactions in a shared latent space, and decode future trajectories in parallel Jia et al. (2023); Liu et al. (2024); Shi et al. (2024). This makes them an appealing foundation for efficient planning. Indeed, recent one-shot planners such as PlanTF and PLUTO have adopted prediction-inspired architectures to generate ego trajectories in a single forward pass Cheng et al. (2024b;a). However, their performance still often lags behind strong iterative or optimization-heavy methods in challenging interactive scenarios.

We argue that this gap does not mainly stem from insufficient model capacity or from the one-shot inference style itself. Instead, it arises because planning requires *decision-aware* interaction reasoning, whereas trajectory prediction is designed for *behavior forecasting*. In many existing one-shot planners, much of this burden is shifted to the planning decoder, which must simultaneously infer which surrounding agents matter, reason about how they may affect the ego vehicle, and generate the final ego trajectory. As scene complexity grows, this leads to increasingly specialized and decoder-heavy planning modules, while still lacking an explicit notion of which interactions are safety-critical and which agents are truly relevant to the ego decision.

This observation suggests a different path. Rather than designing a new planning-specific decoder or relying on iterative generation, we ask whether trajectory prediction backbones can be *directly repurposed* for planning by injecting the missing decision-aware signals. The key obstacle is that prediction models do not explicitly model *safety-critical relevance*: they do not indicate which surrounding agents are most important for ego planning, nor do they explicitly encode how risky future interactions may become.

Safety-critical relevance is not limited to imminent collisions. Courtesy yielding, cut-ins, merging, unprotected left turns, and four-way-stop negotiations all alter relative motion and the evolution of potential trajectory conflicts over the planning horizon. We model perceived safety as a temporal interaction signal whose future evolution supervises the latent representation. In this way, the safety branch captures how potential conflict emerges, persists, and is resolved over time, and associates these dynamics with the motion and map context used for planning.

To address this challenge, we propose SCRP, a prediction-guided trajectory planning framework for safety-critical relevance planning with adaptive agent selection (as detailed in Figure 1(b)). SCRP builds upon a generic *state encoder–interaction encoder–trajectory decoder* architecture inherited

from trajectory prediction models, allowing planning to be performed within a unified one-shot framework. On top of this backbone, SCRP introduces two lightweight yet effective components for decision-aware interaction reasoning. First, a perceived safety modeling mechanism injects explicit future-risk supervision through horizon-wise TTC/TET/TIT prediction. Second, our *risk-conditioned adaptive selection* mechanism converts motion–map–safety relevance into a normalized relevance mass and chooses a scene-dependent cardinality: the smallest retained set whose residual relevance mass falls below a risk- and ambiguity-conditioned tolerance. Thus, sparse scenes can use fewer interaction tokens, while dense, uncertain, or high-risk scenes automatically retain more. By jointly modeling evolving safety and adaptive interaction relevance, SCRP guides the prediction backbone to focus computation where it matters and turns efficient trajectory prediction models into strong trajectory planners.

Our contributions can be summarized as follows:

- **A new perspective on efficient planning through repurposed prediction.** We argue that the key limitation of existing one-shot planners is not the one-shot paradigm itself, but the lack of explicit modeling of safety-critical relevance in multi-agent interaction reasoning. This perspective motivates repurposing trajectory prediction models for planning instead of designing increasingly complex planning-specific modules.
- **A general framework for safety-critical relevance planning.** We introduce SCRP, a general framework that repurposes trajectory prediction backbones for planning without modifying their core architecture. SCRP equips the backbone with two complementary mechanisms: perceived safety modeling, which captures the temporal evolution of future interaction risk, and risk-conditioned adaptive selection, which chooses a scene-specific number of decision-relevant agents from an explicit residual relevance-mass criterion. We further provide a bounded-omission analysis showing how the adaptive selector controls the deviation from full-agent interaction reasoning.
- **Bridging efficiency and planning performance.** Extensive experiments on large-scale autonomous driving benchmarks show that SCRP achieves competitive or superior planning performance while maintaining high efficiency. In closed-loop evaluation, SCRP improves the closed-loop score by 11.99 and 9.62 points on Val14-NR and Val14-R, respectively, and by 9.17 and 8.98 points on Test14-NR and Test14-R, respectively. On nuPlan open-loop evaluation, SCRP reduces minADE, minFDE, and MR by 32.4%, 40.5%, and 26.4% on average across the three prediction backbones. Adding SCRP reduces inference latency by 4.5%–9.1% relative to those backbones while preserving comparable or superior planning performance.

2 Related Work

2.1 Trajectory Planning for Autonomous Driving

Trajectory planning focuses on generating safe and feasible ego motions under complex multi-agent interactions, differing primarily in uncertainty modeling and safety enforcement. Rule-based planners rely on handcrafted heuristics and explicit constraints to ensure interpretability and safety. Classical models such as the Intelligent Driver Model (IDM) Treiber et al. (2000) capture longitudinal interactions but lack lateral maneuvering capabilities, while extensions such as PDM-Closed Dauner et al. (2023) enumerate predefined candidate trajectories and select solutions via rule-based scoring. Although effective in structured environments, these methods are constrained by fixed maneuver templates and struggle to generalize to rare or highly coupled interaction scenarios.

Imitation learning formulates planning as behavior cloning from expert demonstrations, with recent methods enhancing interaction modeling through transformer-based architectures and game-theoretic reasoning, including PlanTF Cheng et al. (2024b), Pluto Cheng et al. (2024a), and GameFormer Huang et al. (2023a). However, their performance critically depends on the coverage of expert data and often degrades under distribution shift, particularly in safety-critical or long-tail situations. Reinforcement learning learns planning policies through direct interaction with the environment by optimizing long-term objectives, thereby enabling adaptation beyond expert data Tang et al. (2026); Huang et al. (2026). Representative approaches such as CarPlanner Zhang et al. (2025a) show strong empirical

performance on large-scale benchmarks, but RL-based planners face challenges in reward design, incur high training costs, and provide limited interpretability in terms of safety reasoning.

Generative approaches, especially diffusion-based planners, model complex multi-modal future distributions through iterative sampling, as exemplified by Diffusion-ES Yang et al. (2024), DiffusionDrive Liao et al. (2025a), and Diffusion Planner Zheng et al. (2025). Despite their expressiveness, the reliance on repeated generation steps and large model architectures often leads to high inference latency, limiting their practicality for real-time planning systems.

2.2 Safety Awareness and Interaction Relevance

Safety awareness is central to robust autonomous driving. Traffic safety metrics, such as Time-to-Collision and its variants, are widely used to quantify collision risk Minderhoud & Bovy (2001) and are commonly incorporated as auxiliary constraints, costs, or post-processing components in prediction and planning Ren et al. (2021); Li et al. (2024); Lu et al. (2025). Recent work further models evolving risk through adaptive risk fields, cognitive safety representations, and learnable integrated planning costs Liao et al. (2025c;d); Huang et al. (2023b). SCRP follows a representation-learning perspective: future safety trajectories provide temporal supervision for interaction features and subsequently guide scene-dependent agent routing. Interaction filtering methods aim to reduce computation by selecting agents based on proximity or learned relevance Shi et al. (2021); Rowe et al. (2023), but they typically use a fixed cardinality or an implicit attention sparsity pattern. In contrast, our selector determines the cardinality from the residual safety-relevance mass of each scene and provides an explicit approximation bound.

3 Methodology

3.1 Philosophy of Our Framework

In this study, we aim to repurpose existing trajectory prediction models for ego trajectory planning without introducing a planning-specific backbone. Despite substantial architectural diversity, modern prediction models largely follow a common *state encoder–interaction encoder–trajectory decoder* pipeline: historical agent states and map elements are embedded, their interactions are modeled in a shared latent space, and future trajectories are decoded in one forward pass Jia et al. (2023); Liu et al. (2024); Shi et al. (2024). This common structure provides a strong and efficient foundation for planning, but prediction and planning differ in what information the interaction model should emphasize. Prediction distributes modeling capacity across all agents, whereas planning only needs to preserve the interactions that can materially affect the ego decision.

SCRP centers the planning adaptation on *safety-critical relevance* while retaining the backbone architecture. It introduces two lightweight mechanisms before the original interaction encoder. Perceived safety modeling injects future supervision describing how ego–agent risk evolves over the planning horizon. Risk-conditioned adaptive selection then uses this representation together with motion and map context to determine which agents should participate in ego-centric interaction reasoning and, crucially, how many of them are needed for the current scene. Both mechanisms are external to the backbone-specific encoder and decoder, which makes the framework applicable to heterogeneous prediction architectures while preserving their one-shot inference procedure.

3.2 Overall Architecture

As illustrated in Figure 2, adapting prediction to planning requires addressing two complementary issues. First, the planner must reason about potential interactions that have not yet materialized in the observed history, such as an approaching merge, cut-in, or yielding conflict. Second, it should avoid spending equal interaction capacity on nearby agents whose future influence on the ego decision is negligible. SCRP addresses the first issue by learning a safety-aware representation from future interaction outcomes, and the second by routing a scene-dependent subset of agents to the interaction encoder.

Concretely, the original state encoder produces agent and map features. The safety branch processes historical motion and safety states and predicts their future evolution, yielding a fused safety

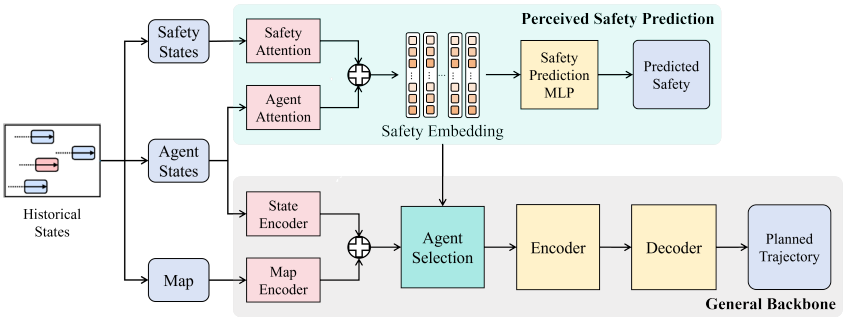


Figure 2: **Overview of the proposed SCRP framework**. SCRP inherits a generic state encoder–interaction encoder–trajectory decoder backbone from trajectory prediction. A perceived safety prediction branch learns future-supervised safety representations, while risk-conditioned adaptive selection converts joint motion–map–safety relevance into a scene-specific interaction cardinality. The retained agent tokens are then processed by the original interaction encoder and trajectory decoder.

representation $\mathbf{S}^{\oplus}$. The adaptive selector combines $\mathbf{S}^{\oplus}$ with motion–map features to construct a safety-relevance distribution over surrounding agents and selects the smallest ranked prefix that satisfies a scene-conditioned residual criterion. The selected agent tokens, together with the map tokens used by the original model, are passed to the unchanged interaction encoder and trajectory decoder. Thus, SCRP changes *what information reaches the planner* rather than replacing how the backbone performs interaction modeling or trajectory decoding.

3.3 Perceived Safety Prediction

To reason about potential future interactions, we leverage perceived safety, which has been shown effective for characterizing driving risk and behavior adaptation Liao et al. (2025b); Lu et al. (2025). We quantify safety with Time-to-Collision (TTC), Time Exposed Time-to-Collision (TET), and Time Integrated Time-to-Collision (TIT) Minderhoud & Bovy (2001). TTC describes prospective conflict under the current relative motion, whereas TET and TIT characterize how long and how strongly the conflict persists over time. Their detailed definitions are provided in Appendix A.1.

A cut-in, merge, unprotected turn, or yielding maneuver changes relative position and velocity continuously; the corresponding TTC/TET/TIT sequence therefore describes how risk emerges, persists, and is resolved over the planning horizon. Future ground-truth safety states are used as auxiliary supervision, shaping the representation with information about prospective interaction outcomes. The resulting embedding provides a temporal safety context that complements historical motion and map geometry during planning.

For historical agent states $\mathcal{S}^{(\mathrm{g})} \in \mathbb{R}^{N_a \times T_h \times C_a}$ and safety states $\mathcal{S}^{(\mathrm{l})} \in \mathbb{R}^{N_a \times T_h \times 3}$, two parallel MLP–attention branches compress temporal information and capture contextual dependencies. We then fuse the two streams and predict the future safety sequence:

$$\mathbf{S}^{(\mathrm{l})} = \phi^{(\mathrm{l})}_{MHA}\Big(\phi^{(\mathrm{l})}_{MLP}(\mathcal{S}^{(\mathrm{l})})\Big), \quad \mathbf{S}^{(\mathrm{g})} = \phi^{(\mathrm{g})}_{MHA}\Big(\phi^{(\mathrm{g})}_{MLP}(\mathcal{S}^{(\mathrm{g})})\Big), \tag{1}$$

$$\mathbf{S}^{\oplus} = \phi^{\oplus}_{MLP}\Big(\mathbf{S}^{(\mathrm{l})} \oplus \mathbf{S}^{(\mathrm{g})}\Big), \qquad \hat{\mathcal{Y}}^{(\mathrm{s})} = \phi^{(\mathrm{s})}_{MLP}(\mathbf{S}^{\oplus}), \tag{2}$$

where $\mathbf{S}^{\oplus} \in \mathbb{R}^{N_a \times F}$ is the fused safety-aware embedding and $\hat{\mathcal{Y}}^{(\mathrm{s})} \in \mathbb{R}^{N_a \times T_f \times 3}$ predicts the three safety channels over the planning horizon T_f. The fused embedding conditions the planning backbone, while the predicted sequence supplies a lightweight horizon-risk signal for adaptive routing. The safety branch is trained with an MSE loss $\mathcal{L}_S$ over future steps and metrics; further implementation details are given in Appendix A.2.

3.4 Risk-Conditioned Adaptive Critical Agent Selection

Trajectory planning is inherently ego-centric, and the interaction-set size can vary substantially across scenes. A sparse highway segment may be explained by only a few agents, whereas a dense merge or intersection can involve several simultaneously relevant participants. We model this variation with

risk-conditioned adaptive selection (RCAS), which determines a scene-specific cardinality from the residual safety-relevance mass.

Following the prediction backbone, agent and map encoders produce $\mathbf{H} \in \mathbb{R}^{N_a \times F}$ and $\mathbf{M} \in \mathbb{R}^{N_m \times F}$. Map-conditioned motion features are fused with the safety representation to obtain a comprehensive agent feature, from which we compute an ego–agent distance:

$$\mathbf{H}^{(\mathrm{r})} = \phi^{(\mathrm{r})}_{MLP}(\mathbf{H} \oplus \mathbf{M}), \qquad \mathbf{H}^{(\mathrm{c})} = \mathbf{H}^{(\mathrm{r})} \oplus \mathbf{S}^{\oplus}, \qquad c_{ij} = \|\mathbf{h}^{(\mathrm{c})}_i - \mathbf{h}^{(\mathrm{c})}_j\|_2. \tag{3}$$

A small c_{ij} indicates strong representation-level relevance. In parallel, the predicted safety sequence is pooled into a horizon risk score $\rho_{ij} \in [0, 1]$ that emphasizes high-risk future moments; its exact construction is given in Appendix A.3. Motion–map similarity and future risk jointly define the normalized safety-relevance mass

$$p_{ij} = \mathrm{softmax}_{j \neq i}(-c_{ij}/\tau_c + \beta \rho_{ij}) \,. \tag{4}$$

This distribution provides both an ordering of agents and a measure of how concentrated the scene relevance is. Let $r_i = \sum_{j \neq i} p_{ij} \rho_{ij}$ be aggregate future risk and let $h_i \in [0, 1]$ be the normalized entropy of $\{p_{ij}\}$, with larger h_i indicating that relevance is spread across more agents. We define

$$\begin{aligned} \epsilon_i &= \epsilon_{\min} + (\epsilon_{\max} - \epsilon_{\min})(1 - r_i)(1 - h_i), \\ K_i &= \min\left\{ k \in [K_{\min}, K_{\max}] : \sum_{q=1}^{k} p_{i(q)} \geq 1 - \epsilon_i \right\}, \end{aligned} \tag{5}$$

where $p_{i(1)} \geq \cdots \geq p_{i(N_a - 1)}$ are the sorted relevance masses. High-risk or diffuse scenes receive a tighter residual budget and therefore retain more agents, while concentrated low-risk scenes use fewer interaction tokens. If the target cannot be reached before the cap, $K_i = K_{\max}$. The ego token is always retained, and the discrete routing decision is treated as stop-gradient; no additional cardinality loss is introduced.

The selector couples two complementary notions of scene complexity. The entropy term reflects *interaction multiplicity*: when relevance is distributed across several agents, more tokens are required to represent the scene faithfully. The aggregate-risk term reflects *interaction consequence*: even a relatively concentrated scene receives additional capacity when its predicted future conflicts are severe. Because both quantities are computed from normalized relevance mass, the routing decision depends on the structure of the current interaction pattern rather than the raw number of detections. This is useful in crowded scenes containing many distant or weakly coupled agents, where density alone is a poor proxy for planning relevance.

Proposition 3.1 (Adaptive bounded omission)**.** *Let $\bar{\epsilon}_i$ be the realized relevance mass omitted by RCAS and let η_i bound the ℓ_1 error between estimated relevance and the unknown normalized full-agent influence. If interaction values are bounded by B and the downstream interaction–decoder map is L_D-Lipschitz, then $\|\hat{\mathbf{Y}}^{\mathrm{full}}_i - \hat{\mathbf{Y}}^{\mathrm{RCAS}}_i\|_2 \leq 2BL_D(\bar{\epsilon}_i + \eta_i)$. Whenever the residual target is feasible, Eq.* (5) *also yields the minimum-cardinality ranked subset satisfying that target.*

The proposition connects the adaptive interaction budget to a controlled approximation of full-agent reasoning: smaller residual mass yields a tighter trajectory-deviation bound, while RCAS uses the smallest ranked subset that meets the scene-specific criterion. The full risk-pooling definition, assumptions, proof, and induced planning-cost bound are provided in Appendices A.3 and B.

3.5 Trajectory Planning and Model Training

After adaptive selection, the retained safety-aware agent features and original map tokens are passed through the backbone interaction encoder and trajectory decoder, preserving one-shot decoding. Training jointly optimizes the backbone trajectory objective $\mathcal{L}_T$ and future-safety objective $\mathcal{L}_S$, so safety supervision shapes both the planning representation and the routing signal while trajectory supervision remains backbone-specific. The exact joint objective is provided in Appendix A.4.

Table 1: Closed-loop planning results on the nuPlan dataset. denotes the best result in each case. NR denotes non-reactive mode, and R denotes reactive mode.

Method	Val14						Test14					
	NR			R			NR			R		
	Score↑	Comf.↑	TTC↑	Score↑	Comf.↑	TTC↑	Score↑	Comf.↑	TTC↑	Score↑	Comf.↑	TTC↑
BeTop	81.00	84.85	80.47	82.94	86.94	88.73	84.34	92.28	89.91	83.76	93.47	88.45
SCRP+BeTop	**93.69**	**97.14**	**92.86**	**93.06**	**98.10**	**92.38**	**92.43**	**94.78**	**91.91**	**91.36**	**95.81**	**92.29**
MTR	80.00	75.30	**90.40**	82.60	79.70	**91.20**	80.00	82.60	88.40	76.00	78.30	86.70
SCRP+MTR	**91.30**	**89.71**	89.22	**90.80**	**87.38**	90.78	**91.02**	**85.29**	**91.01**	**89.73**	**85.27**	**89.96**
EDA	78.60	**81.10**	81.10	82.50	80.10	83.20	81.10	78.10	82.60	82.30	80.70	84.50
SCRP+EDA	**91.45**	79.40	**93.47**	**92.37**	**85.12**	**93.02**	**93.61**	**88.67**	**94.39**	**89.57**	**82.30**	**89.76**
PlanTF	75.86	**93.80**	80.10	77.55	**94.90**	84.70	85.44	**94.90**	85.43	84.11	**97.10**	87.32
SCRP+PlanTF	**86.98**	86.54	**93.27**	**87.83**	88.07	**97.25**	**90.49**	90.37	**90.14**	**91.41**	95.29	**92.92**
Diffusion-ES	90.31	87.36	89.42	88.23	85.20	90.57	87.36	85.19	86.54	86.62	83.17	86.59
Diffusion Planner	90.22	88.10	88.69	83.17	82.58	88.91	88.82	87.38	89.41	84.86	83.25	85.94
SAH-Drive	89.64	87.80	91.32	88.42	88.17	90.23	86.92	88.15	89.01	87.63	86.30	90.24

Table 2: Open-loop planning results on the WOMD and nuPlan datasets. For each baseline and its SCRP-enhanced counterpart, the better value for each metric is highlighted in bold.

Method	WOMD			nuPlan				
	minADE↓	minFDE↓	MR↓	minADE↓	minFDE↓	MR↓	Score↑	
							Val14	Test14
BeTop	0.7181	1.4957	0.1806	1.5390	2.8657	0.3085	83.24	81.94
SCRP+BeTop	**0.6245**	**1.2229**	**0.1423**	**0.8995**	**1.3236**	**0.2188**	**90.26**	**89.72**
MTR	0.7451	1.5678	0.2017	1.4211	2.826	0.3111	80.19	79.57
SCRP+MTR	**0.6967**	**1.4224**	**0.1861**	**1.0865**	**1.8210**	**0.2917**	**87.58**	**86.48**
EDA	0.7036	1.4646	0.1724	1.0131	1.9934	0.2972	81.35	80.45
SCRP+EDA	**0.6140**	**1.2027**	**0.1346**	**0.6887**	**1.3516**	**0.1667**	**90.14**	**89.32**
Diffusion-ES	0.7942	1.5763	0.1904	0.7527	1.3884	0.1831	88.42	86.35
Diffusion Planner	0.7023	1.4950	0.1826	0.6905	1.3683	0.1810	86.69	87.31
SAH-Drive	0.7164	1.5033	0.1798	0.6942	1.3527	0.1698	87.94	88.30

4 EXPERIMENTS

4.1 EXPERIMENTAL SETUP

We evaluate SCRP on two real-world datasets, WOMD Ettinger et al. (2021) and nuPlan Caesar et al. (2021), under both open-loop and closed-loop planning protocols. Open-loop performance is measured by minimum average displacement error (minADE), minimum final displacement error (minFDE), and miss rate (MR). Closed-loop evaluation uses the Val14 and Test14 benchmarks in non-reactive and reactive settings and reports the overall Score, comfort score (Comf.), and time-to-collision (TTC). Dataset statistics, observation and prediction horizons, benchmark hardware, baseline configurations, and hyperparameter settings are provided in Appendices C-E.

4.2 MAIN RESULTS

Closed-loop Planning. To evaluate SCRP across prediction architectures, we integrate it into three representative trajectory prediction models, MTR Shi et al. (2024), BeTop Liu et al. (2024), and EDA Lin et al. (2024). These models share a Transformer-based encoder but differ substantially in their attention-based decoding mechanisms, allowing us to assess whether SCRP provides consistent benefits across decoder designs. We also compare the SCRP-enhanced models with three SOTA trajectory planning models, Diffusion-ES Yang et al. (2024), Diffusion Planner Zheng et al. (2025), and SAH-Drive Fan et al. (2025). We begin with closed-loop planning on nuPlan Val14 and Test14 under non-reactive (NR) and reactive (R) settings. For this evaluation, we additionally integrate

Table 3: Closed-loop ablation results on the nuPlan dataset. We report the overall closed-loop Score under non-reactive (NR) and reactive (R) settings on Val14 and Test14. Complete SCRP is compared with the variant that removes the perceived safety prediction branch. The best results are highlighted.

Method	Val14		Test14	
	NR	R	NR	R
SCRP+BeTop	**93.69**	**93.06**	**92.43**	**91.36**
w/o Safety	88.36	89.11	86.93	86.62
SCRP+MTR	**91.30**	**90.80**	**91.02**	**89.73**
w/o Safety	87.01	87.68	86.46	85.92
SCRP+EDA	**91.45**	**92.37**	**93.61**	**89.57**
w/o Safety	88.73	89.71	90.71	86.93
SCRP+PlanTF	**86.98**	**87.83**	**90.49**	**91.41**
w/o Safety	84.71	84.62	86.74	88.42

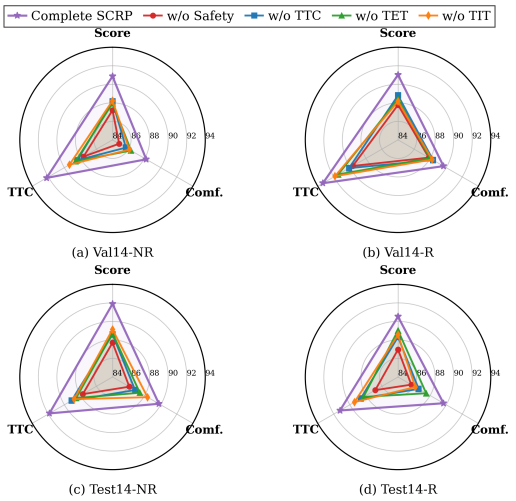


Figure 3: Ablation of safety components on nuPlan. Each radar plot reports values averaged over four SCRP-enhanced methods and compares complete SCRP with variants that remove the full safety branch or single safety metric.

SCRP with PlanTF Cheng et al. (2024b), which supports closed-loop execution. The results are shown in Table 1.

Across all evaluated backbones, SCRP consistently improves closed-loop Score in all four settings. Averaged over BeTop, MTR, EDA, and PlanTF, SCRP brings gains of 11.99, 9.62, 9.17, and 8.98 points on Val14-NR, Val14-R, Test14-NR, and Test14-R, respectively. The improvements are consistent across individual backbones as well, with gains ranging from 5.05 to 13.73 points. These results show that SCRP strengthens long-horizon interaction reasoning and safety-aware decision making in complex closed-loop environments.

SCRP-enhanced planners are also highly competitive with dedicated planning architectures. SCRP achieves the best Score in all four settings, outperforming the strongest dedicated planning baseline by 3.38, 4.64, 4.79, and 3.78 points on Val14-NR, Val14-R, Test14-NR, and Test14-R, respectively. SCRP+PlanTF also obtains the highest TTC score in the Val14-R and Test14-R settings, indicating improved safety under reactive interactions. Although SCRP does not improve Comfort for every backbone, especially for PlanTF, it remains competitive on comfort while delivering a stronger overall Score and safety performance. For PlanTF, Comfort does not improve after adding SCRP because the original PlanTF often stops near surrounding agents, producing conservative low-variation motions that are favored by the Comfort metric but are less desirable for interactive driving. These results demonstrate that SCRP enables lightweight prediction-based planners to match or surpass specialized planning architectures while retaining a general and modular design. In addition to Score, Comfort, and TTC, we further report speed limit compliance in Appendix F, since respecting speed constraints is another important requirement for safe and rule-compliant closed-loop planning. We also report the scenario-level standard deviation of the closed-loop Score in Appendix G.

Open-loop Planning. We next assess the same three prediction backbones in the open-loop setting on WOMD and nuPlan, with results shown in Table 2. On WOMD, SCRP improves minADE by **10.8%** on average, with reductions of *13.0%*, *6.5%*, and *12.8%* for BeTop, MTR, and EDA, respectively. Consistent gains in minFDE and MR further indicate improved accuracy and fewer failure cases across different prediction architectures. On nuPlan, SCRP reduces minADE, minFDE, and MR by **32.4%**, **40.5%**, and **26.4%** on average across the three backbones, respectively, demonstrating consistent improvements in open-loop planning quality. Notably, on WOMD, SCRP+EDA and SCRP+BeTop outperform all three dedicated planning models, including Diffusion-ES, Diffusion Planner, and SAH-Drive, across both minADE and MR. On nuPlan, SCRP consistently improves all evaluated backbones, yielding average Score gains of **7.73** points on Val14 and **7.85** points on Test14. Moreover, SCRP+BeTop achieves the highest Score on both benchmarks, while remaining

competitive on the other key metrics. These results demonstrate that safety-critical relevance effectively strengthens the planning capability of prediction backbones, enabling them to match or surpass specialized planning architectures while retaining a lightweight, model-agnostic inference pipeline.

4.3 Ablation Studies

We evaluate the role of perceived safety prediction in closed-loop planning by removing the entire safety branch while keeping the remaining architecture unchanged, denoted as *w/o Safety*. As shown in Table 3, removing this branch consistently lowers the closed-loop score on nuPlan across all backbones and evaluation settings. For example, SCRP+BeTop drops from 93.69 to 88.36 on Val14-NR and from 92.43 to 86.93 on Test14-NR. The same trend appears for MTR, EDA, and PlanTF, where complete SCRP outperforms w/o Safety in all 16 Val14/Test14 and NR/R comparisons. This confirms that perceived safety prediction provides essential information for planning.

Figure 3 further studies the contribution of individual safety metrics by removing TTC, TET, or TIT from the perceived safety input. Each value in the radar plots is averaged over the four evaluated SCRP-enhanced backbones, namely BeTop, MTR, EDA, and PlanTF. Complete SCRP achieves the strongest radar profile across Score, Comfort, and TTC in all four settings, indicating that these metrics provide complementary signals. Removing any one of them weakens the safety representation.

Notably, the improvement is not limited to TTC. Complete SCRP also improves Comfort compared with the safety-ablated variants, suggesting that safety-aware relevance does not simply make the planner more conservative. Instead, it helps identify decision-relevant agents and generate smoother ego motions. This observation is consistent with the use of TTC, TET, and TIT as temporal interaction descriptors: maneuvers involving yielding, merging, or waiting change not only instantaneous collision risk but also the duration and severity of prospective conflicts. These results support the core design of SCRP: *future safety supervision and multi-metric safety modeling jointly improve closed-loop safety and planning quality.*

RCAS constitutes the second core component of SCRP. We therefore evaluate its scene-dependent cardinality separately from the safety ablations above by comparing RCAS with fixed interaction budgets and analyzing how the retained K_i changes with scene density, predicted future risk, and maneuver type; the complete analysis is provided in Appendix H.

4.4 Efficiency Analyses

Adding SCRP reduces average per-scenario latency by **4.5%–9.1%** across all prediction-based backbones. RCAS retains approximately ten agents on average in the most interactive nuPlan Val14-R setting, while using smaller interaction sets in simpler scenes and larger sets when relevance is diffuse or future risk is high. This allocation matches the non-uniform demand of interaction reasoning: routine lane following is often dominated by few nearby influences, whereas merges and intersections require simultaneous reasoning over several participants. Dedicated diffusion and hybrid planners require approximately **10**$\times$–**60**$\times$ more inference time. Runtime and training-convergence results are provided in Appendices I–J, while the adaptive-cardinality behavior of RCAS is analyzed in Appendix H; qualitative results are provided in Appendix K.

5 Conclusion

In this paper, we propose SCRP, a safety-aware trajectory planning framework that repurposes trajectory prediction models for interactive driving. Perceived safety modeling injects future interaction information into the latent representation, while risk-conditioned adaptive selection allocates a scene-specific interaction set through a residual relevance-mass criterion. The bounded-omission analysis relates this routing decision to the fidelity of full-agent interaction reasoning. Across WOMD and nuPlan, SCRP consistently improves open-loop accuracy, closed-loop planning quality, and safety-related metrics while reducing inference latency. These results support safety-critical relevance as a practical interface for transferring efficient one-shot prediction backbones to planning. This balance of planning quality and efficiency is particularly valuable in dense, rapidly changing traffic scenes, where safe decisions must be made promptly.

AI USE STATEMENT

Generative AI tools were used for language polishing and LaTeX assistance. All scientific claims, experimental results, mathematical arguments, and the final manuscript were reviewed and verified by the authors.

REFERENCES

Holger Caesar, Juraj Kabzan, Kok Seang Tan, Whye Kit Fong, Eric Wolff, Alex Lang, Luke Fletcher, Oscar Beijbom, and Sammy Omari. nuplan: A closed-loop ml-based planning benchmark for autonomous vehicles. *arXiv preprint arXiv:2106.11810*, 2021.

Jie Cheng, Yingbing Chen, and Qifeng Chen. PLUTO: Pushing the limit of imitation learning-based planning for autonomous driving, 2024a.

Jie Cheng, Yingbing Chen, Xiaodong Mei, Bowen Yang, Bo Li, and Ming Liu. Rethinking imitation-based planners for autonomous driving. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 14123–14130. IEEE, 2024b.

Daniel Dauner, Marcel Hallgarten, Andreas Geiger, and Kashyap Chitta. Parting with misconceptions about learning-based vehicle motion planning. In *7th Annual Conference on Robot Learning*, 2023.

Scott Ettinger, Shuyang Cheng, Benjamin Caine, Chenxi Liu, Hang Zhao, Sabeek Pradhan, Yuning Chai, Ben Sapp, Charles R Qi, Yin Zhou, et al. Large scale interactive motion forecasting for autonomous driving: The waymo open motion dataset. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 9710–9719, 2021.

Yuqi Fan, Zhiyong Cui, Zhenning Li, Yilong Ren, and Haiyang Yu. SAH-drive: A scenario-aware hybrid planner for closed-loop vehicle trajectory generation. In *Forty-Second International Conference on Machine Learning*, 2025.

Yihan Hu, Jiazhi Yang, Li Chen, Keyu Li, Chonghao Sima, Xizhou Zhu, Siqi Chai, Senyao Du, Tianwei Lin, Wenhai Wang, et al. Planning-oriented autonomous driving. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 17853–17862, 2023.

Jun Huang, Beining Wu, Qiang Duan, Liang Dong, and Shui Yu. A fast uav trajectory planning framework in ris-assisted communication systems with accelerated learning via multithreading and federating. *IEEE Transactions on Mobile Computing*, 2025.

Zhao Huang, Bowen Zhang, Zhongzhu Li, and Di Lin. Occdriver: Future occupancy guided dual-branch trajectory planner in autonomous driving. In *The Fourteenth International Conference on Learning Representations*, 2026.

Zhiyu Huang, Haochen Liu, and Chen Lv. Gameformer: Game-theoretic modeling and learning of transformer-based interactive prediction and planning for autonomous driving. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3903–3913, 2023a.

Zhiyu Huang, Haochen Liu, Jingda Wu, and Chen Lv. Differentiable integrated motion prediction and planning with learnable cost function for autonomous driving. *IEEE transactions on neural networks and learning systems*, 35(11):15222–15236, 2023b.

Xiaosong Jia, Penghao Wu, Li Chen, Yu Liu, Hongyang Li, and Junchi Yan. Hdgt: Heterogeneous driving graph transformer for multi-agent trajectory prediction via scene encoding. *IEEE transactions on pattern analysis and machine intelligence*, 45(11):13860–13875, 2023.

Boqian Li, Shengbo Wang, Zhenyuan Guo, Song Zhu, Junjian Huang, Junwei Sun, Guanghui Wen, and Shiping Wen. Safe control framework of multi-agent systems from a performance enhancement perspective. *IEEE Transactions on Automation Science and Engineering*, 2024.

Yue Li, Meng Tian, Dechang Zhu, Jiangtong Zhu, Zhenyu Lin, Zhiwei Xiong, and Xinhai Zhao. Drive-r1: Bridging reasoning and planning in vlms for autonomous driving with reinforcement learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 6708–6716, 2026.

Bencheng Liao, Shaoyu Chen, Haoran Yin, Bo Jiang, Cheng Wang, Sixu Yan, Xinbang Zhang, Xiangyu Li, Ying Zhang, Qian Zhang, and Xinggang Wang. DiffusionDrive: Truncated diffusion model for end-to-end autonomous driving. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12037–12047, 2025a.

Haicheng Liao, Zhenning Li, Guohui Zhang, Keqiang Li, and Chengzhong Xu. Toward human-like trajectory prediction for autonomous driving: A behavior-centric approach. *Transportation Science*, 2025b.

Haicheng Liao, Zhenning Li, Kaiqun Zhu, Keqiang Li, and Chengzhong Xu. Sa-tp {2}: A safety-aware trajectory prediction and planning model for autonomous driving. *IEEE Transactions on Robotics*, 2025c.

Haicheng Liao, Chengyue Wang, Kaiqun Zhu, Yilong Ren, Bolin Gao, Shengbo Eben Li, Chengzhong Xu, and Zhenning Li. Minds on the move: Decoding trajectory prediction in autonomous driving with cognitive insights. *IEEE Transactions on Intelligent Transportation Systems*, 26(5):6101–6115, 2025d.

Longzhong Lin, Xuewu Lin, Tianwei Lin, Lichao Huang, Rong Xiong, and Yue Wang. Eda: Evolving and distinct anchors for multimodal motion prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 3432–3440, 2024.

Haochen Liu, Li Chen, Yu Qiao, Chen Lv, and Hongyang Li. Reasoning multi-agent behavioral topology for interactive autonomous driving. *Advances in Neural Information Processing Systems*, 37:92605–92637, 2024.

Haochen Liu, Tianyu Li, Haohan Yang, Li Chen, Caojun Wang, Ke Guo, Haochen Tian, Hongchen Li, Hongyang Li, and Chen Lv. Reinforced refinement with self-aware expansion for end-to-end autonomous driving. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2026.

Yuhuan Lu, Zhen Zhang, Rufan Bai, Han Liu, and Wei Wang. Dual-view interaction-aware lane change prediction for autonomous driving. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 28222–28230, 2025.

Yuhuan Lu, Pengpeng Xu, Wei Wang, Zhen Zhang, Han Liu, and Xiping Hu. Knowlcp: Knowledge augmented lane change prediction for autonomous driving. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 38979–38987, 2026.

Michiel M Minderhoud and Piet HL Bovy. Extended time-to-collision measures for road traffic safety assessment. *Accident Analysis & Prevention*, 33(1):89–97, 2001.

Xuanchi Ren, Tao Yang, Li Erran Li, Alexandre Alahi, and Qifeng Chen. Safety-aware motion prediction with unseen vehicles for autonomous driving. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 15731–15740, 2021.

Luke Rowe, Martin Ethier, Eli-Henry Dykhne, and Krzysztof Czarnecki. Fjmp: Factorized joint multi-agent motion prediction over learned directed acyclic interaction graphs. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 13745–13755, 2023.

Liushuai Shi, Le Wang, Chengjiang Long, Sanping Zhou, Mo Zhou, Zhenxing Niu, and Gang Hua. Sgcn: Sparse graph convolution network for pedestrian trajectory prediction. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 8994–9003, 2021.

Shaoshuai Shi, Li Jiang, Dengxin Dai, and Bernt Schiele. Mtr++: Multi-agent motion prediction with symmetric scene modeling and guided intention querying. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(5):3955–3971, 2024.

Xiaolong Tang, Meina Kan, Shiguang Shan, and Xilin Chen. Plan-r1: Safe and feasible trajectory planning as language modeling. In *The Fourteenth International Conference on Learning Representations*, 2026.

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

Brian Yang, Huangyuan Su, Nikolaos Gkanatsios, Tsung-Wei Ke, Ayush Jain, Jeff Schneider, and Katerina Fragkiadaki. Diffusion-es: Gradient-free planning with diffusion for autonomous and instruction-guided driving. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 15342–15353, 2024.

Dongkun Zhang, Jiaming Liang, Ke Guo, Sha Lu, Qi Wang, Rong Xiong, Zhenwei Miao, and Yue Wang. Carplanner: Consistent auto-regressive trajectory planning for large-scale reinforcement learning in autonomous driving. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 17239–17248, 2025a.

Pin Zhang, Ke Lin, Duantengchuan Li, Zixun Fu, Yuefeng Cai, Bing Li, Huan Yu, and Ming Li. Daplanner: Dual-agent framework with multi-modal large language model for autonomous driving motion planning. *Applied Soft Computing*, pp. 113625, 2025b.

Yinan Zheng, Ruiming Liang, Kexin ZHENG, Jinliang Zheng, Liyuan Mao, Jianxiong Li, Weihao Gu, Rui Ai, Shengbo Eben Li, Xianyuan Zhan, and Jingjing Liu. Diffusion-based planning for autonomous driving with flexible guidance. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=wM2sfVgMDH.

Yupeng Zheng, Zebin Xing, Qichao Zhang, Bu Jin, Pengfei Li, Yuhang Zheng, Zhongpu Xia, Kun Zhan, Xianpeng Lang, Yaran Chen, et al. Planagent: A multi-modal large language agent for closed-loop vehicle motion planning. *arXiv preprint arXiv:2406.01587*, 2024.

A Details of Methodology

A.1 Details of Safety Metrics

This section provides the detailed definitions of the safety metrics used in Section 3.3.

Time-to-Collision (TTC). TTC represents the remaining time before two agents collide if they continue along their current trajectories. Since conventional TTC is defined for one-dimensional longitudinal motion, we extend it to a two-dimensional setting. The TTC of the ego agent i with respect to a surrounding agent j at time step t is:

$$e_{TTC_{i,j}^t} = -\frac{d_{i,j}^t}{\hat{d}_{i,j}^t}, \tag{6}$$

where $d_{i,j}^t$ denotes the Euclidean distance between agents i and j, and $\hat{d}_{i,j}^t$ denotes the temporal rate of change of this distance:

$$\begin{cases} d_{i,j}^t = \sqrt{(p_i^t - p_j^t)^\top (p_i^t - p_j^t)}, \\ \hat{d}_{i,j}^t = \frac{1}{d_{i,j}^t}(p_i^t - p_j^t)^\top (v_i^t - v_j^t). \end{cases} \tag{7}$$

Here, (p_i^t, v_i^t) and (p_j^t, v_j^t) denote the two-dimensional positions and velocities of agents i and j, respectively.

Time Exposed Time-to-Collision (TET). While TTC reflects instantaneous risk, TET measures the cumulative duration during which TTC remains within a safety-critical range over a temporal horizon t_H:

$$e_{TET_i^t} = \sum_{t_k=t-t_H+1}^{t} \delta_i(t_k) \cdot \tau_{\text{sc}}, \tag{8}$$

where τ_{sc} is set to 0.1 s. The indicator function is:

$$\delta_i(t_k) = \begin{cases} 1, & 0 \le e_{TTC_i^{t_k}} \le e_{TTC^*}, \\ 0, & \text{otherwise}, \end{cases} \tag{9}$$

where e_{TTC^*} is the safety-critical TTC threshold, set to 2.5 s following Lu et al. (2026).

Time Integrated Time-to-Collision (TIT). TIT further distinguishes different risk severities by integrating the TTC deficit over time:

$$e_{TIT_i^t} = \sum_{t_k=t-t_H+1}^{t} \left[e_{TTC^*} - e_{TTC_i^{t_k}} \right] \cdot \tau_{\text{sc}}, \quad \forall\, 0 \le e_{TTC_i^{t_k}} \le e_{TTC^*}. \tag{10}$$

Risk-oriented TTC. A larger raw TTC indicates a safer interaction, whereas TET and TIT increase with risk. To keep the three safety channels in a consistent direction, we convert TTC into a risk-oriented quantity:

$$\tilde{e}_{TTC_{i,j}^t} = \begin{cases} e_{TTC^*} - e_{TTC_{i,j}^t}, & 0 \le e_{TTC_{i,j}^t} \le e_{TTC^*}, \\ 0, & \text{otherwise}. \end{cases} \tag{11}$$

After this conversion, larger values of $\tilde{e}_{TTC}$, e_{TET}, and e_{TIT} consistently indicate higher collision risk and lower perceived safety.

A.2 Additional Details on Perceived Safety Prediction

As discussed in Section 3.3, the perceived safety branch predicts safety states over the same horizon as trajectory decoding. Therefore, it does not produce only a one-step warning, but models the evolution of interaction risk throughout the planning horizon.

The predicted safety sequence $\hat{\mathcal{Y}}^{(\mathrm{s})}$ serves as an auxiliary supervised target and a lightweight routing signal for the adaptive selector. The fused embedding $\mathbf{S}^{\oplus}$ conditions the backbone, while the normalized predicted sequence is pooled to determine the interaction cardinality. Thus, the safety metrics enter SCRP through temporal representation learning and interaction routing.

The safety loss $\mathcal{L}_S$ is computed as an MSE loss over all future steps and all three safety indicators. Before adaptive selection, each safety channel is normalized using statistics computed on the training split, and the risk-oriented TTC transformation in Eq. (11) ensures that all three channels share the same risk direction.

A.3 Additional Details on Risk-Conditioned Adaptive Selection

RCAS is applied before interaction encoding. Let $\mathbf{H} \in \mathbb{R}^{N_a \times F}$ and $\mathbf{M} \in \mathbb{R}^{N_m \times F}$ be agent and map features from the prediction backbone. We form

$$\mathbf{H}^{(\mathrm{r})} = \phi^{(\mathrm{r})}_{MLP}(\mathbf{H} \oplus \mathbf{M}), \qquad \mathbf{H}^{(\mathrm{c})} = \mathbf{H}^{(\mathrm{r})} \oplus \mathbf{S}^{\oplus}, \tag{12}$$

$$c_{ij} = \|\mathbf{h}^{(\mathrm{c})}_i - \mathbf{h}^{(\mathrm{c})}_j\|_2. \tag{13}$$

The distance term measures compatibility in the joint motion–map–safety representation.

Horizon risk pooling. Let $\bar{y}^{(s)}_{jtm} \in [0,1]$ be the normalized predicted risk of agent j at future step t for metric $m \in \{\mathrm{TTC}, \mathrm{TET}, \mathrm{TIT}\}$, with larger values indicating higher risk. We use soft maximum pooling,

$$a_{jtm} = \frac{\exp(\bar{y}^{(s)}_{jtm}/\tau_s)}{\sum_{u=1}^{T_f}\sum_{n=1}^{3}\exp(\bar{y}^{(s)}_{jun}/\tau_s)}, \qquad \rho_{ij} = \sum_{t=1}^{T_f}\sum_{m=1}^{3} a_{jtm}\bar{y}^{(s)}_{jtm}. \tag{14}$$

This pooling emphasizes hazardous future moments while retaining information from the complete horizon.

Scene risk and ambiguity. Using Eq. (4), aggregate risk and normalized relevance entropy are

$$r_i = \sum_{j \neq i} p_{ij}\rho_{ij}, \qquad h_i = -\frac{1}{\log(N_a - 1)}\sum_{j \neq i} p_{ij}\log(p_{ij} + 10^{-8}), \tag{15}$$

with $h_i = 0$ for a single surrounding candidate. Since $r_i, h_i \in [0,1]$, Eq. (5) yields a residual budget in $[\epsilon_{\min}, \epsilon_{\max}]$. The retained set is the prefix of the sorted masses with cardinality K_i, and the ego token is always retained. If the target mass is not reached before $K_{\max}$, the selector returns $K_{\max}$ and the theory uses the realized residual $\bar{\epsilon}_i$.

We use $K_{\min} = 4$, $K_{\max} = 20$, $\epsilon_{\min} = 0.05$, $\epsilon_{\max} = 0.25$, $\tau_c = 1$, $\tau_s = 0.2$, and $\beta = 1$ across datasets and backbones. $K_{\max}$ provides a deterministic worst-case interaction budget, whereas the realized K_i varies scene by scene.

A.4 Additional Details on Trajectory Planning and Model Training

After adaptive critical agent selection, the retained comprehensive state features are passed to the interaction encoder and trajectory decoder of the chosen prediction backbone. Since these features contain the fused safety embedding $\mathbf{S}^{\oplus}$, trajectory decoding is conditioned on safety-aware context learned from future safety sequences over the same horizon T_f. This allows SCRP to reuse existing prediction architectures while steering ego trajectory generation toward safer future interactions.

We retain the original trajectory loss of each baseline model, denoted as $\mathcal{L}_T$, to preserve compatibility with different prediction backbones. Meanwhile, the perceived safety module is trained with a

supervised safety loss $\mathcal{L}_S$, which encourages the learned representation to encode interaction-relevant future safety information. The overall objective is:

$$\mathcal{L} = \lambda \mathcal{L}_S + (1-\lambda)\mathcal{L}_T, \tag{16}$$

where $\lambda \in [0,1]$ balances safety prediction and trajectory planning, and is selected empirically.

The design of SCRP is further supported by theoretical analysis in Appendix B, where we quantify how future-safety prediction error propagates into the risk score and prove a cardinality-minimal, bounded-omission guarantee for the adaptive selector.

B Theoretical Analysis of Future-Risk Supervision and Adaptive Selection

B.1 Notation and Setup

For a scene centered on ego agent i, let $\mathcal{A}_i$ be the set of $N_a - 1$ surrounding agents. The adaptive selector produces a normalized relevance vector $\mathbf{p}_i = (p_{ij})_{j\in\mathcal{A}_i}$ from Eq. (4), a selected subset $\mathcal{S}_i \subseteq \mathcal{A}_i$, and an omitted set $\mathcal{O}_i = \mathcal{A}_i \setminus \mathcal{S}_i$. We denote its estimated residual mass by

$$\bar{\epsilon}_i = \sum_{j\in\mathcal{O}_i} p_{ij} = 1 - \sum_{j\in\mathcal{S}_i} p_{ij}. \tag{17}$$

Let $\mathbf{q}_i = (q_{ij})$ denote an idealized normalized influence distribution that measures the true contribution of surrounding agents to ego planning, with $q_{ij} \geq 0$ and $\sum_j q_{ij} = 1$. We quantify relevance-estimation error by $\eta_i = \|\mathbf{p}_i - \mathbf{q}_i\|_1$.

B.2 Future-Safety Prediction Controls Risk-Score Error

The selector uses a temporal pooling functional $g(\cdot)$ to map the normalized predicted safety sequence of each agent to ρ_{ij}. Because softmax pooling is smooth on the compact cube $[0,1]^{3T_f}$, it is Lipschitz; let L_g denote a valid Lipschitz constant under the Euclidean norm.

Lemma B.1 (Safety-sequence error to risk-score error)**.** *Let $\widehat{\mathbf{y}}_j^{(s)}$ and $\mathbf{y}_j^{(s)}$ be the predicted and ground-truth normalized future safety sequences for agent j. Then*

$$|\hat{\rho}_{ij} - \rho^\star_{ij}| \leq L_g \left\| \widehat{\mathbf{y}}_j^{(s)} - \mathbf{y}_j^{(s)} \right\|_2. \tag{18}$$

Consequently, if the safety head has mean-squared error at most σ_s^2, then $\mathbb{E}|\hat{\rho}_{ij} - \rho^\star_{ij}| \leq L_g \sigma_s$.

Proof. The first inequality follows directly from Lipschitz continuity of g. For the expectation, apply Jensen's inequality followed by Cauchy–Schwarz:

$$\mathbb{E}|\hat{\rho}_{ij} - \rho^\star_{ij}| \leq L_g\, \mathbb{E}\|\widehat{\mathbf{y}}_j^{(s)} - \mathbf{y}_j^{(s)}\|_2 \leq L_g \sqrt{\mathbb{E}\|\widehat{\mathbf{y}}_j^{(s)} - \mathbf{y}_j^{(s)}\|_2^2} \leq L_g \sigma_s.$$

□

This lemma formalizes why horizon-wise safety supervision is useful for routing: improved prediction of the evolving TTC/TET/TIT sequence directly tightens the error of the scalar risk signal that enters adaptive selection.

B.3 Minimal-Cardinality Residual-Mass Selection

Proposition B.2 (Cardinality minimality)**.** *Let $p_{i(1)} \geq \cdots \geq p_{i(N_a-1)}$ and suppose the target tolerance ϵ is feasible for some $k \leq K_{\max}$. Define*

$$K^\star = \min\left\{ k \in [K_{\min}, K_{\max}] : \sum_{q=1}^{k} p_{i(q)} \geq 1 - \epsilon \right\}.$$

Then the top-$K^\star$ prefix has minimum cardinality among all admissible subsets $\mathcal{S} \subseteq \mathcal{A}_i$ with $K_{\min} \leq |\mathcal{S}| \leq K_{\max}$ and $\sum_{j\notin\mathcal{S}} p_{ij} \leq \epsilon$.

Proof. For any subset of size k, the sum of its elements is at most the sum of the k largest masses, $\sum_{q=1}^{k} p_{i(q)}$. Therefore, if the top-k prefix fails to retain mass $1-\epsilon$, no other admissible size-k subset can satisfy the residual constraint. By construction, the top-$K^\star$ prefix is the first admissible size that satisfies it, proving minimality. □

B.4 Bounded Planning Deviation

We next connect omitted relevance mass to the final planned trajectory. Consider the full-agent interaction aggregation for the ego token

$$\mathbf{z}_i^{\text{full}} = \sum_{j \in \mathcal{A}_i} q_{ij} \mathbf{v}_{ij}, \tag{19}$$

where $q_{ij} \geq 0$, $\sum_j q_{ij} = 1$, and $\|\mathbf{v}_{ij}\|_2 \leq B$. This covers the standard normalized attention aggregation used by the evaluated Transformer-style backbones; for multi-layer encoders, the subsequent layers are absorbed into the downstream Lipschitz map below. Given selected set $\mathcal{S}_i$, the retained influence is renormalized as

$$\mathbf{z}_i^{\text{sel}} = \sum_{j \in \mathcal{S}_i} \frac{q_{ij}}{1 - r_i^{\text{omit}}} \mathbf{v}_{ij}, \qquad r_i^{\text{omit}} = \sum_{j \in \mathcal{O}_i} q_{ij}. \tag{20}$$

Lemma B.3 (Perturbation of renormalized interaction aggregation)**.** *If $r_i^{\text{omit}} < 1$ and $\|\mathbf{v}_{ij}\|_2 \leq B$, then*

$$\left\|\mathbf{z}_i^{\text{full}} - \mathbf{z}_i^{\text{sel}}\right\|_2 \leq 2B r_i^{\text{omit}}. \tag{21}$$

Proof. Write $r = r_i^{\text{omit}}$. Since $\sum_{j \in \mathcal{S}_i} q_{ij} = 1 - r$,

$$\mathbf{z}_i^{\text{full}} - \mathbf{z}_i^{\text{sel}} = \sum_{j \in \mathcal{O}_i} q_{ij} \mathbf{v}_{ij} - \frac{r}{1-r} \sum_{j \in \mathcal{S}_i} q_{ij} \mathbf{v}_{ij}. \tag{22}$$

Applying the triangle inequality and $\|\mathbf{v}_{ij}\|_2 \leq B$ gives

$$\left\|\mathbf{z}_i^{\text{full}} - \mathbf{z}_i^{\text{sel}}\right\|_2 \leq Br + \frac{r}{1-r} B(1-r) = 2Br.$$

□

Assumption B.4 (Downstream Lipschitz decoding)**.** Let G denote the remaining interaction layers and trajectory decoder after the aggregation above. There exists $L_D < \infty$ such that

$$\|G(\mathbf{z}) - G(\mathbf{z}')\|_2 \leq L_D \|\mathbf{z} - \mathbf{z}'\|_2. \tag{23}$$

For finite-weight neural networks with Lipschitz activations, such a constant is bounded by the product of layer operator norms.

Theorem B.5 (Adaptive bounded-omission guarantee)**.** *Under Assumption B.4, let $\mathcal{S}_i$ be the subset returned by RCAS and let $\bar{\epsilon}_i$ be its actual residual mass under $\mathbf{p}_i$. If $\|\mathbf{p}_i - \mathbf{q}_i\|_1 \leq \eta_i$, then*

$$\left\|\hat{\mathbf{Y}}_i(\mathcal{A}_i) - \hat{\mathbf{Y}}_i(\mathcal{S}_i)\right\|_2 \leq 2BL_D \left(\bar{\epsilon}_i + \eta_i\right). \tag{24}$$

When Eq. (5) *reaches its target before the hard cap, $\bar{\epsilon}_i \leq \epsilon_i$ and hence the right-hand side is at most $2BL_D(\epsilon_i + \eta_i)$.*

Proof. For the omitted set $\mathcal{O}_i$,

$$r_i^{\text{omit}} = \sum_{j \in \mathcal{O}_i} q_{ij} \leq \sum_{j \in \mathcal{O}_i} p_{ij} + \sum_{j \in \mathcal{O}_i} |q_{ij} - p_{ij}| \tag{25}$$

$$\leq \bar{\epsilon}_i + \|\mathbf{q}_i - \mathbf{p}_i\|_1 \tag{26}$$

$$\leq \bar{\epsilon}_i + \eta_i. \tag{27}$$

Lemma B.3 yields an aggregation perturbation at most $2B(\bar{\epsilon}_i + \eta_i)$. Applying Assumption B.4 proves Eq. (24). □

Corollary B.6 (Planning-cost bound)**.** *If a planning objective $J(\mathbf{Y})$ is L_J-Lipschitz in the ego trajectory, then*

$$\left|J\left(\hat{\mathbf{Y}}_i(\mathcal{A}_i)\right) - J\left(\hat{\mathbf{Y}}_i(\mathcal{S}_i)\right)\right| \leq 2BL_J L_D \left(\bar{\epsilon}_i + \eta_i\right). \tag{28}$$

Proof. Apply the Lipschitz property of J to Theorem B.5. □

Table 4: Shared hyperparameters of the adaptive SCRP selector. The same cardinality-control parameters are used for all datasets and backbones. $\bar{K}$ reports the observed average retained cardinality.

Setting	$K_{\min}$	$K_{\max}$	$\epsilon_{\min}$	$\epsilon_{\max}$	τ_c	τ_s	β	λ	$\bar{K}$	**Selection**
WOMD open-loop	4	20	0.05	0.25	1.0	0.2	1.0	0.5	8.74	Adaptive
nuPlan open-loop	4	20	0.05	0.25	1.0	0.2	1.0	0.5	9.21	Adaptive
nuPlan Val14-NR	4	20	0.05	0.25	1.0	0.2	1.0	0.5	9.56	Adaptive
nuPlan Val14-R	4	20	0.05	0.25	1.0	0.2	1.0	0.5	10.02	Adaptive
nuPlan Test14-NR	4	20	0.05	0.25	1.0	0.2	1.0	0.5	9.47	Adaptive
nuPlan Test14-R	4	20	0.05	0.25	1.0	0.2	1.0	0.5	9.86	Adaptive

B.5 Why the Cardinality Grows in Difficult Scenes

The scene-conditioned budget in Eq. (5) is monotone in both risk and ambiguity:

$$\frac{\partial \epsilon_i}{\partial r_i} = -(\epsilon_{\max} - \epsilon_{\min})(1 - h_i) \le 0, \qquad \frac{\partial \epsilon_i}{\partial h_i} = -(\epsilon_{\max} - \epsilon_{\min})(1 - r_i) \le 0. \tag{29}$$

Therefore, increasing either future risk r_i or relevance entropy h_i cannot loosen the residual budget. For a fixed ranked mass vector, a smaller tolerance can only preserve or increase the minimum feasible K_i. This provides a direct theoretical explanation for the empirical trend in Appendix H: RCAS allocates more interaction capacity to dense, ambiguous, or safety-critical scenes and less to simple scenes.

B.6 Computational Implication

If the agent-interaction stage uses quadratic self-attention, replacing $N_a - 1$ surrounding-agent tokens by K_i tokens reduces the agent-agent interaction term from $\mathcal{O}(N_a^2 F)$ to $\mathcal{O}((K_i + 1)^2 F)$. Because K_i is scene-dependent, the expected interaction cost scales with $\mathbb{E}[(K_i + 1)^2]$ rather than a worst-case fixed budget. The cap $K_{\max}$ preserves deterministic worst-case latency, while Proposition B.2 ensures that no larger ranked subset is used than required by the residual criterion.

C Additional Experimental Details

The WOMD dataset contains approximately 487,000 training scenes, 44,000 validation scenes, and 44,000 testing scenes. We use 1 second of historical trajectory data to predict 8 seconds into the future. The nuPlan dataset contains approximately one million training cases with 8-second horizons, and we use 2 seconds of historical observations as input. For closed-loop evaluation, all applicable methods are tested under both non-reactive and reactive settings. The overall closed-loop Score is averaged across all scenarios and ranges from 0 to 100, where higher values indicate better planning performance. All experiments are conducted on a common platform with four NVIDIA GeForce RTX 4090 GPUs and Ubuntu 22.04.5. Baseline-specific configurations and the selected SCRP hyperparameters are reported in the following sections.

D Planning Baselines

We briefly introduce the baseline methods integrated into our SCRP framework, as well as representative closed-loop trajectory planning baselines, as follows:

- **BeTop** Liu et al. (2024): a learning-based planner that explicitly reasons about multi-agent interactions via behavioral topology, guiding trajectory generation with topological priors to ensure compliant and socially consistent planning under complex interactive scenarios.
- **MTR** Shi et al. (2024): a transformer-based trajectory planner that generates multimodal future trajectories using learnable intention queries and interaction-aware decoding, enabling scene-compliant planning through joint reasoning over agent dynamics and map context.

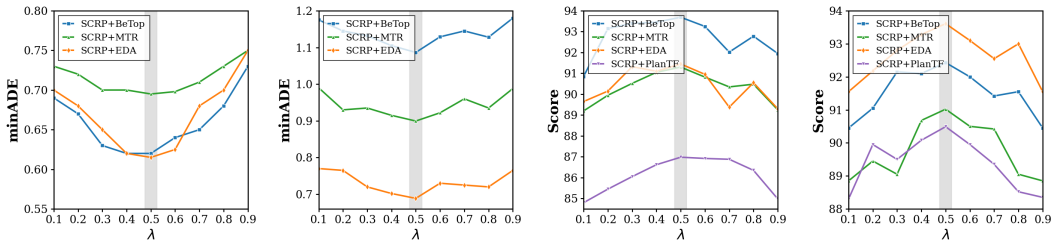


Figure 4: Hyperparameter study of λ in SCRP across different baselines. From left to right, the four panels report WOMD open-loop minADE, nuPlan open-loop minADE, nuPlan Val14 closed-loop Score, and nuPlan Test14 closed-loop Score. Lower minADE is better, while higher Score is better.

- **EDA** Lin et al. (2024): an anchor-based trajectory planner that generates diverse planning candidates by evolving and selecting distinct trajectory anchors, improving multimodal coverage while maintaining efficient and stable trajectory selection for downstream planning.
- **PlanTF** Cheng et al. (2024b): a purely imitation-based planning transformer proposed on the nuPlan benchmark. It studies key design choices for learning-based planners, including input features and data augmentation, and shows that a well-designed imitation-based planner can achieve competitive closed-loop performance without relying on hand-crafted rules or post-optimization.
- **Diffusion-ES** Yang et al. (2024): a learning-based planner that combines a diffusion model with evolutionary search, iteratively refining trajectory samples to obtain high-quality planning solutions.
- **Diffusion Planner** Zheng et al. (2025): a transformer-based diffusion model that produces trajectories without rule-based heuristics by jointly handling prediction and planning, guided by a classifier for high-quality sampling.
- **SAH-Drive** Fan et al. (2025): a scenario-aware hybrid planner that combines a rule-based planner with a diffusion-based learning planner, dynamically switching between them through a dual-timescale decision mechanism to balance efficiency and robustness in closed-loop planning.

E Hyperparameter Settings

We use one selector configuration across all backbones and evaluation protocols: $K_{\min} = 4$, $K_{\max} = 20$, $\epsilon_{\min} = 0.05$, $\epsilon_{\max} = 0.25$, $\tau_c = 1$, $\tau_s = 0.2$, and $\beta = 1$. The task-loss coefficient λ is selected from $\{0.1, 0.2, \ldots, 0.9\}$, with $\lambda = 0.5$ used in the reported experiments (Figure 4). The selector parameters are shared by WOMD, nuPlan open-loop, Val14, and Test14.

F Speed Limit Compliance

Speed limit compliance is an important indicator for closed-loop planning, since a planner must not only avoid collisions and generate comfortable motion, but also follow traffic rules under interactive driving conditions. Although the main closed-loop results in Table 1 report Score, Comfort, and TTC, speed compliance provides an additional view of whether the planned trajectory remains rule-compliant. We therefore include this analysis in the appendix as a complementary evaluation.

Table 5 shows that SCRP consistently improves speed limit compliance over all corresponding base models. For BeTop, SCRP raises the compliance score from 92.35 to 98.94 on Val14-NR, from 96.98 to 98.37 on Val14-R, from 96.80 to 97.36 on Test14-NR, and from 92.54 to 96.21 on Test14-R. The same trend is observed for MTR, EDA, and PlanTF, where SCRP improves the compliance score in every evaluated setting.

The gains are particularly clear for PlanTF. SCRP+PlanTF improves the score by 8.65, 10.68, 5.35, and 12.75 points on Val14-NR, Val14-R, Test14-NR, and Test14-R, respectively. This indicates that safety-aware interaction reasoning does not simply encourage conservative behavior, but also helps the planner generate trajectories that better respect traffic constraints. SCRP+PlanTF achieves the

Table 5: Speed limit compliance results on the nuPlan dataset. denotes the best result in each case. NR denotes non-reactive mode, and R denotes reactive mode.

Method	Val14		Test14	
	NR↑	R↑	NR↑	R↑
BeTop	92.35	96.98	96.80	92.54
SCRP+BeTop	**98.94**	**98.37**	**97.36**	**96.21**
MTR	92.65	89.77	88.54	90.31
SCRP+MTR	**94.42**	**94.80**	**91.28**	**92.57**
EDA	90.30	92.11	89.64	90.69
SCRP+EDA	**92.16**	**94.25**	**91.40**	**91.49**
PlanTF	89.53	88.54	92.41	85.83
SCRP+PlanTF	**98.18**	**99.22**	**97.76**	**98.58**
Diffusion-ES	94.27	92.56	93.89	93.10
Diffusion Planner	93.72	94.51	92.17	92.88
SAH-Drive	93.19	93.59	93.64	92.80

Table 6: Scenario-level closed-loop Score (mean ± standard deviation) on nuPlan. NR and R denote non-reactive and reactive evaluation, respectively.

Method	Val14		Test14	
	NR Score↑	R Score↑	NR Score↑	R Score↑
Diffusion-ES	90.31 ± 14.62	88.23 ± 16.18	87.36 ± 17.05	86.62 ± 17.84
Diffusion Planner	90.22 ± 14.91	83.17 ± 19.47	88.82 ± 15.88	84.86 ± 18.73
SAH-Drive	89.64 ± 15.27	88.42 ± 16.05	86.92 ± 17.23	87.63 ± 16.54
SCRP+BeTop (Ours)	$\mathbf{93.69 \pm 10.84}$	$\mathbf{93.06 \pm 11.27}$	$\mathbf{92.43 \pm 12.15}$	$\mathbf{91.36 \pm 12.92}$

best compliance score in three of the four settings, while SCRP+BeTop obtains the best result on Val14-NR.

Compared with dedicated planning architectures, SCRP-enhanced models are also highly competitive. SCRP+BeTop and SCRP+PlanTF both exceed the diffusion-based and hybrid planning baselines in most settings, with SCRP+PlanTF reaching 99.22 on Val14-R and 98.58 on Test14-R. These results further support the conclusion from the main closed-loop evaluation: by selecting safety-relevant agents and conditioning planning on perceived safety, SCRP improves not only collision-related safety metrics such as TTC, but also rule-compliant behavior such as speed limit compliance.

G Closed-Loop Score Variability

Table 6 reports the scenario-level mean and standard deviation of the closed-loop Score for SCRP+BeTop and three dedicated planning baselines, with means matching Table 1. SCRP+BeTop achieves both the highest mean Score and the lowest standard deviation in all four settings. Its standard deviation is lower than that of the least variable dedicated baseline by 3.78, 4.78, 3.73, and 3.62 points on Val14-NR, Val14-R, Test14-NR, and Test14-R, respectively, a reduction of approximately 22%–30%. This advantage persists across both validation and test splits and under non-reactive and reactive evaluation, showing that SCRP+BeTop combines strong average performance with more consistent scenario-level Scores.

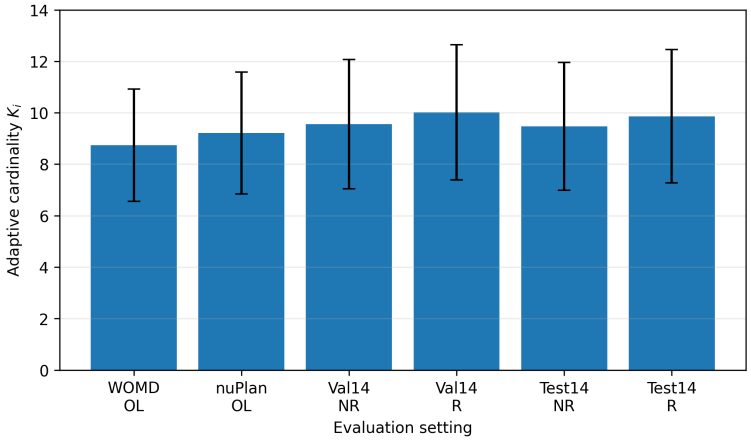


Figure 5: Observed adaptive cardinality across benchmarks. Bars report the mean retained K_i and error bars show one standard deviation. The selector uses a single shared parameterization yet allocates more interaction tokens to the reactive nuPlan settings.

Table 7: Comparison of fixed-cardinality and adaptive agent selection. Each entry is averaged over the SCRP-enhanced backbones available for the corresponding benchmark.

Benchmark	**Metric**	$K{=}6$	$K{=}8$	$K{=}10$	$K{=}12$	**Adaptive**	$\bar{K}$
WOMD open-loop	minADE↓	0.682	0.658	0.653	0.655	**0.645**	8.74
nuPlan open-loop	minADE↓	0.941	0.913	0.902	0.907	**0.892**	9.21
nuPlan Val14-R	Score↑	88.46	90.00	90.65	90.67	**91.02**	10.02
nuPlan Test14-R	Score↑	88.91	89.82	90.09	90.03	**90.52**	9.86

H Analysis of Risk-Conditioned Adaptive Selection

RCAS is the second core component of SCRP and is evaluated independently from hyperparameter selection. We examine whether scene-dependent cardinality improves over fixed interaction budgets and whether the resulting K_i changes systematically with scene complexity.

Table 7 compares RCAS against fixed cardinalities. RCAS achieves the strongest result in each reported setting while retaining approximately ten agents on average in Val14-R. The mean cardinality summarizes a scene-dependent distribution rather than a fixed operating point: Figures 5 and 6 show how K_i varies with scene density, predicted future risk, and maneuver type. Sparse or low-risk scenes typically use fewer agents, whereas dense merges and intersections receive a larger interaction budget.

Table 8 provides a fine-grained comparison over $K \in \{6, \ldots, 14\}$. Fixed-cardinality performance improves as the interaction set grows and then saturates around the same scale as the mean adaptive cardinality. RCAS remains stronger because it allocates interaction capacity per scene rather than imposing one cardinality on both sparse and dense interactions.

The adaptive statistics in Figures 5–6 provide an additional interpretation. The monotonic growth of K_i with local density and future-risk percentile is precisely the behavior expected from the derivatives of Eq. (5) derived in Appendix B. This connects the empirical adaptation pattern to the theoretical residual-risk criterion rather than treating K_i as an opaque learned gate.

I Inference Efficiency Details

Figure 7 reports the average inference time per test scenario under the same evaluation setup. Across both datasets, integrating SCRP reduces inference time for all prediction-based baselines by 4.5%–

Table 8: Fine-grained fixed-K comparison on nuPlan Val14-R. Adaptive values are exactly the SCRP results reported in Table 1; all fixed-cardinality variants use the same backbone and training protocol.

Method	K=6	K=7	K=8	K=9	K=10	K=11	K=12	K=13	K=14	**Adaptive**
SCRP+BeTop	89.17	90.44	92.12	92.48	92.71	92.73	92.69	92.62	92.55	**93.06**
SCRP+MTR	87.96	88.15	89.33	89.62	90.32	90.44	90.41	90.35	90.28	**90.80**
SCRP+EDA	90.64	90.67	91.53	92.14	92.09	92.11	92.06	92.03	91.98	**92.37**
SCRP+PlanTF	86.07	86.15	87.01	87.52	87.49	87.55	87.51	87.47	87.42	**87.83**

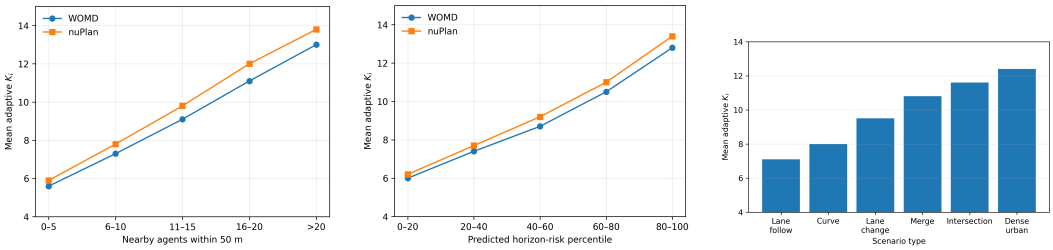


Figure 6: How adaptive cardinality changes with scene properties. Left: K_i increases with the number of nearby agents. Middle: K_i increases with the percentile of predicted horizon risk. Right: more interactive maneuver classes require larger sets. These trends match Eq. (5): higher risk or more diffuse relevance tightens the residual budget.

9.1%. The smaller reduction for EDA is explained by the larger share of computation spent on anchor-based trajectory generation in that backbone, which is unaffected by agent filtering.

Dedicated planning models such as SAH-Drive and diffusion-based planners require approximately 10×–70× more inference time than the SCRP-enhanced one-shot systems. This comparison reflects differences between complete planning architectures and is not attributed solely to the SCRP modules. Together with the effectiveness results, the measurements show that risk-conditioned adaptive selection reduces redundant interaction computation without sacrificing trajectory quality.

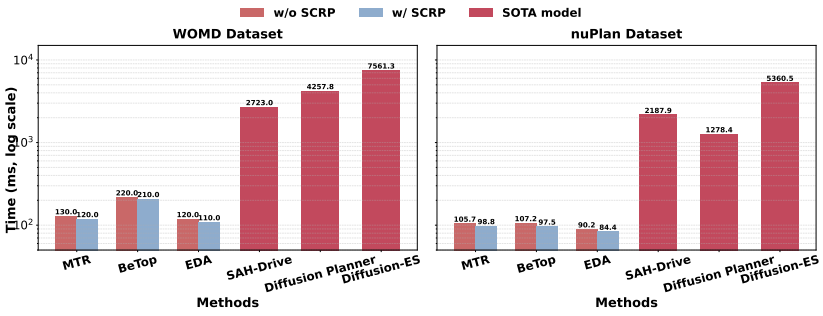
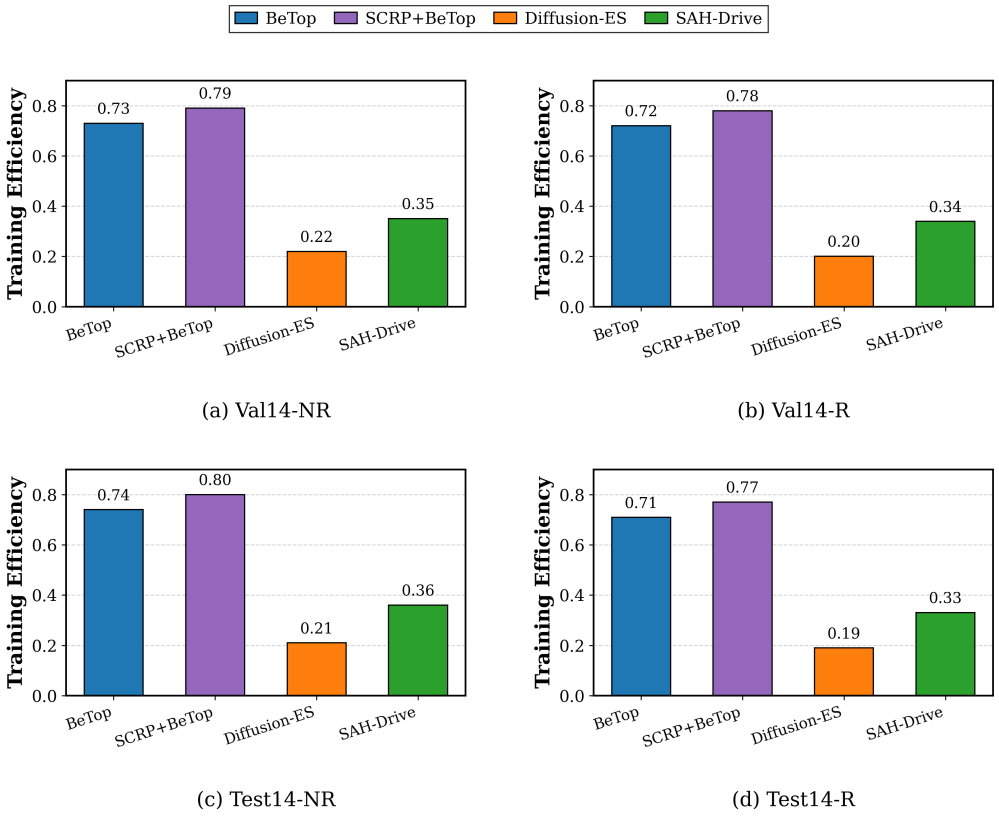


Figure 7: Inference time comparison on the WOMD (left) and nuPlan (right) datasets. Within each prediction backbone, adding SCRP reduces latency; iterative planning models are included as a separate cross-architecture reference.

J Training Efficiency

We further evaluate the training efficiency of SCRP on the nuPlan benchmark. In addition to inference speed, training convergence is also important for practical planning systems, since planning models are often retrained or fine-tuned when new data, scenarios, or safety requirements are introduced. To quantify convergence behavior, we define a simple training efficiency metric based on the training loss curve:

$$\mathrm{Eff}_{\mathrm{train}} = \frac{\mathcal{L}_{10}}{\mathcal{L}_{\mathrm{last}}}, \tag{30}$$



(a) Val14-NR (b) Val14-R

(c) Test14-NR (d) Test14-R

Figure 8: Training efficiency comparison on four nuPlan evaluation settings. We report $\mathrm{Eff}_{\mathrm{train}} = \mathcal{L}_{10}/\mathcal{L}_{\mathrm{last}}$, where larger values indicate more effective convergence after the early training stage. SCRP+BeTop is compared with the original BeTop backbone and two diffusion-based planning baselines, Diffusion-ES and SAH-Drive.

where $\mathcal{L}_{10}$ denotes the training loss at the 10-th epoch and $\mathcal{L}_{\mathrm{last}}$ denotes the training loss at the final epoch. This metric measures how much the loss decreases after the early training stage. A larger value indicates faster and more effective convergence.

Figure 8 reports the results on Val14-NR, Val14-R, Test14-NR, and Test14-R. We compare our best-performing SCRP-enhanced model, SCRP+BeTop, with the original BeTop backbone and two strong diffusion-based planning baselines, Diffusion-ES and SAH-Drive. BeTop and SCRP+BeTop both show substantially higher training efficiency than the diffusion-based planners across all four settings. This result suggests that repurposing a trajectory prediction backbone for planning leads to a more tractable optimization problem than training heavy generative planners for direct planning. The prediction backbone already provides structured interaction modeling and trajectory decoding, so SCRP can inject safety-critical relevance without introducing the optimization burden of iterative diffusion-style generation.

SCRP+BeTop also shows favorable training efficiency relative to the original BeTop backbone. Perceived safety prediction provides additional supervision, while risk-conditioned adaptive selection reduces weakly relevant interaction tokens, together yielding a more focused optimization trajectory for the repurposed planning model.

Together, these results show that lightweight safety-aware reasoning can be integrated with the efficient optimization structure of trajectory prediction backbones while improving planning performance.

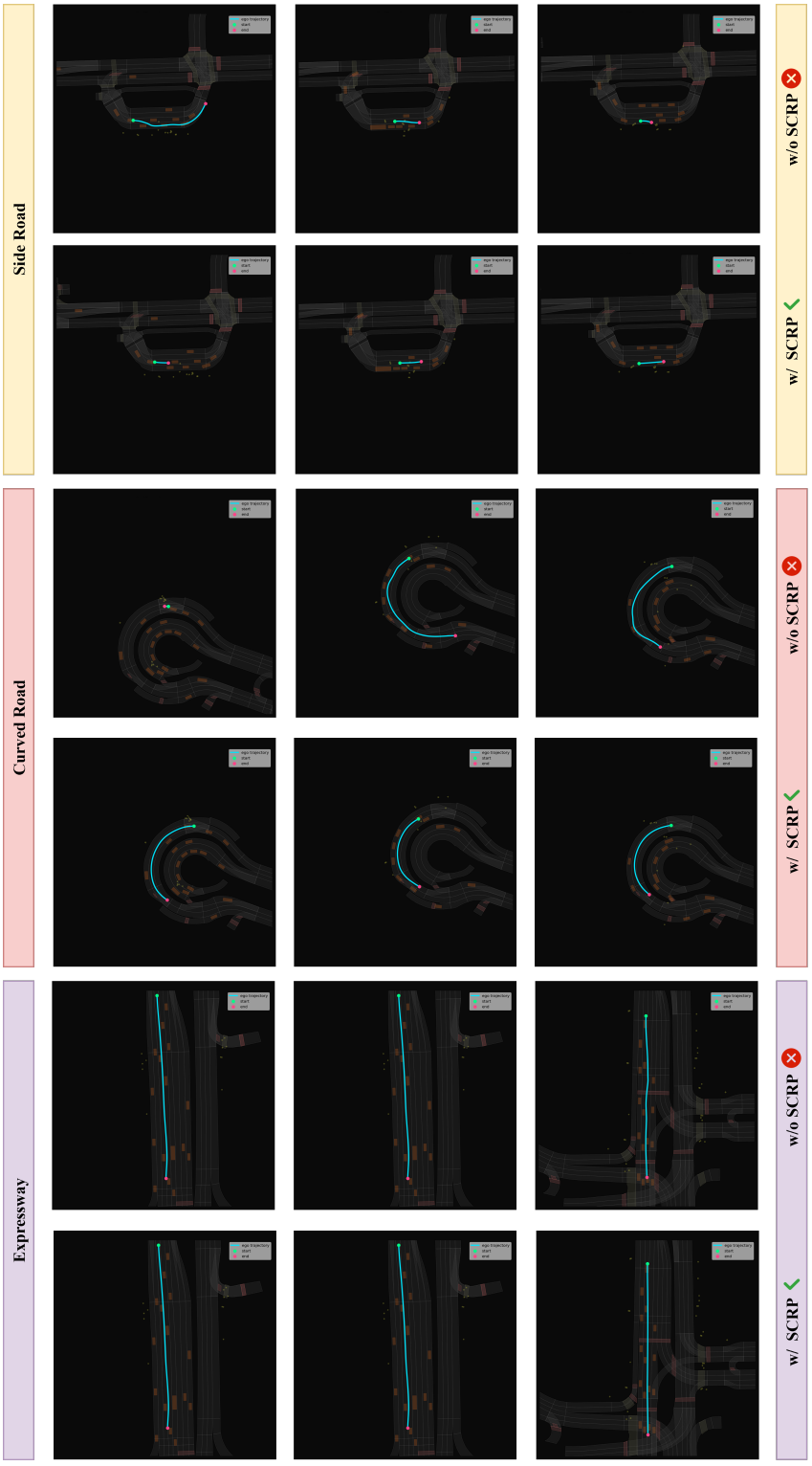


Figure 9: Qualitative comparison between the original PlanTF planner and SCRP-enhanced PlanTF on the nuPlan dataset. The three groups correspond to different road structures: side road, curved road, and expressway. In each group, the upper row shows PlanTF without SCRP, while the lower row shows SCRP+PlanTF. Without SCRP, the planner may collide with surrounding agents, change lanes without sufficient safety consideration, stop unnecessarily, or move toward occupied regions. With SCRP, the planner focuses on safety-critical interactions and produces safer and more stable trajectories.

K Additional Qualitative Results

Figure 9 provides additional qualitative results on the nuPlan dataset. We compare the original PlanTF planner, denoted as w/o SCRP in the figure, with SCRP+PlanTF, denoted as w/ SCRP. The examples are organized into three road structures: side road, curved road, and expressway. Each group contains paired results before and after adding SCRP.

In the side-road cases, the original PlanTF planner shows unstable and unsafe behaviors. In the left example, the planned trajectory collides with a nearby agent. In the middle example, the planner does not sufficiently consider safety-critical interactions and performs an abrupt lane change. In the right example, the planner stops in place, indicating that it fails to make a confident decision under interaction uncertainty. After adding SCRP, the planner avoids these failure modes and generates more reasonable trajectories by focusing on agents that are relevant to the ego vehicle's future motion.

In the curved-road cases, the baseline is strongly affected by surrounding traffic. In the left example, PlanTF stops after observing a vehicle ahead, resulting in overly conservative behavior. In the middle and right examples, it continues to move toward regions occupied by other vehicles, showing insufficient awareness of future interaction risk. SCRP+PlanTF produces safer trajectories in these cases, maintaining progress while avoiding unsafe proximity to surrounding agents.

In the expressway cases, the original planner may change lanes without adequate safety consideration and may even drive close to or across lane boundaries. These behaviors suggest that the planner lacks an explicit mechanism for identifying which agents and lane-level interactions are critical for the ego decision. With SCRP, the generated trajectories become more stable and lane-consistent, indicating better interaction filtering and safety-aware planning.

Across these examples, perceived safety modeling and risk-conditioned adaptive selection concentrate interaction reasoning on agents that influence the ego vehicle's future motion. The resulting trajectories are safer and more decisive across different road structures, complementing the quantitative gains in Score, TTC, speed limit compliance, and other closed-loop metrics.

Limitations. Although these visualizations show clear improvements in representative nuPlan scenarios, they cover only a limited set of road structures and interaction patterns. SCRP may still face challenges in rare traffic configurations, noisy perception inputs, or cases where safety-critical agents are difficult to identify from motion and safety cues alone. A more systematic failure-case analysis across broader scenario categories is left for future work.