OC-RAP: OBSERVATION-CONSISTENT RECOVERY-AFFORDANCE PLANNER

Anonymous authors
Paper under double-blind review

ABSTRACT

Safety-critical planning does not end with avoiding initial contact. Near the safety boundary, a planner should preserve recoverable continuations; after contact, it should separate, avoid secondary events, and return to a viable state. Under partial observation, however, branch-wise planning can be optimistic: each latent future may admit a recovery even when observationally indistinguishable futures require incompatible recovery actions. We formalize this mismatch as an *oracle-to-deployable recoverability gap* and introduce **OC-RAP**, the **Observation-Consistent Recovery-Affordance Planner**, as an observation-consistent deployable recovery-planning framework. For each candidate prefix, OC-RAP rolls out actuator-realizable recovery options under latent roots. Its core **Observation-Consistent Measurable Executable Recovery Operator (OC-MERO)** performs lower-tail aggregation of each *fixed* option over roots compatible with the same post-prefix observation before option selection, preventing hidden root identity from authorizing an unrealizable contingent recovery. A signed margin represents positive recovery reserve and negative recovery debt, providing a unified recovery semantics before and after contact. **Recovery-Informed Feasibility and Abstention (RIFA)** first requires candidate-prefix validity and conservative absolute deployable-recovery evidence; only then may the fixed within-admission preference score choose among admitted candidates, and otherwise the planner preserves the exact nominal action. Across frozen real-world-motion cohorts evaluated in closed loop, OC-RAP achieves zero intervention in Safe while maintaining nominal-utility preservation (NUP) $= 1$. On 250 Near-Contact scenes, it obtains the best non-collision clearance 5th percentile (p05; 0.661 m), time-to-collision (TTC) p05 (0.130 s), and TTC-deficit area under the curve (AUC; 0.630 s^2), although the best comparator has a lower collision rate and another has shorter critical-TTC exposure. On Contact, OC-RAP attains the shortest overlap, lowest penetration AUC, and largest terminal clearance on the fully paired 8-anchor physical comparison; a complementary continuous re-contact audit reports the highest initial-separation rate (88.5%) and the lowest post-separation re-contact rate (25.0%) and exposure (8.6%). Matched ablations support the roles of observation consistency, executable recovery geometry, absolute admission, and exact-nominal abstention.

1 INTRODUCTION

Collision avoidance is necessary for autonomous driving, but it is not a complete description of safety near the boundary of failure. Before contact, a planner should preserve useful braking, evasive, or stabilization continuations rather than optimize only nominal short-horizon quality. After contact, planning remains safety-critical: the vehicle must stabilize, separate in finite time, avoid re-contact and secondary collisions, and return to a viable route state (Koopman, 2024; Wang et al., 2023; Parseh et al., 2023; Wang et al., 2024; Yang et al., 2025; Ghosh et al., 2026). This motivates treating *recoverability* as a property of a candidate action rather than treating collision merely as a terminal label.

The main difficulty is informational. Suppose several plausible latent futures remain indistinguishable after the ego vehicle executes a candidate prefix. A branch-wise oracle may select one recovery maneuver in one future and another maneuver in a different future. The deployed vehicle cannot make that choice until its observations distinguish the futures. Thus, "every latent branch has a recovery" is

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.

weaker than "the candidate has a recovery the deployed vehicle can select." This gap becomes acute near contact, where small hidden differences in another agent's response, in friction, or in actuation delay can change which continuation is useful.

OC-RAP makes this information pattern explicit. Each candidate prefix is paired with a finite library of actuator-realizable recovery continuations. For roots that remain compatible with the same post-prefix observation, **OC-MERO** evaluates the lower tail of each fixed recovery option *before* allowing option selection. In the hard-partition limit, the selected option is therefore observation-measurable; the practical implementation uses a soft compatibility relaxation. This is a recovery-specific non-anticipativity construction: the novelty is not the general principle that decisions may depend only on available information (Shapiro et al., 2021; Kaelbling et al., 1998), but an executable lower-tail operator that applies it to learned latent futures and recovery selection.

The same signed recovery quantity is used on both sides of contact. A positive margin represents recovery *reserve* to preserve or enlarge; a negative margin represents recovery *debt* to repay. This removes the need for separate Near-Contact and Contact experts. OC-RAP also separates recovery evidence from intervention preference. **RIFA** first enforces candidate-prefix validity and absolute deployable-recovery admission; preference scoring is evaluated only within the admitted set, and the exact nominal action has priority when intervention is unsupported. This design matters because recovery existence and deployability do not by themselves establish that an intervention is preferable to continuing nominal behavior.

Our contributions are:

1. **Observation-consistent deployable recoverability.** We formalize the oracle-to-deployable gap induced by partial observation and introduce OC-MERO, which performs lower-tail aggregation over observation-compatible roots for each fixed recovery option before option selection. We prove policy-class oracle dominance from the information restriction and observation measurability of the OC-MERO choice in the hard-partition limit.
2. **Executable signed recovery with role-isolated intervention.** Recovery options are projected into actuator-feasible commands and evaluated by a non-compensatory signed margin, yielding a unified reserve/debt semantics before and after contact. RIFA keeps absolute recovery admission upstream of non-nominal preference and fails closed to the exact nominal action.
3. **Frozen closed-loop evaluation across Safe, Near-Contact, and Contact.** We use one recovery model and one selector functional form across all strata, an observation-legal target lock, a common simulator execution path and metric implementation, and no post-lock tuning. The evaluation reports cross-method trade-offs rather than a single aggregate score, and matched inference-time ablations isolate the deployed mechanisms.

2 Related Work

Post-impact and collision-mitigation planning. Research on post-impact planning and control addresses stabilization, controllability, trajectory restoration, secondary-collision avoidance, and active safety after an initial collision (Parseh et al., 2023; Wang et al., 2024; Yang et al., 2025). Representative baselines include integrated post-impact model predictive control (MPC) (Wang et al., 2023), post-impact braking (PIB) (Lu et al., 2017), post-impact motion planning with a time-varying linear–quadratic regulator (TVLQR) (Wang et al., 2022), post-collision trajectory restoration (Ghosh et al., 2026), feedforward–feedback compensation MPC (FCC-MPC) (Cao et al., 2021), and post-impact sliding-mode control (SMC) with fault-tolerant quadratic-program (QP) allocation (Ao et al., 2022); collision-mitigation planners additionally optimize motion when impact may be unavoidable (Li et al., 2022; Bosia et al., 2024). These works establish that safety remains an active control problem after contact. OC-RAP addresses a different question: whether a recovery remains selectable from the information available to the deployed planner. It uses the same signed recovery object before and after contact.

Partial observability, safety filters, and risk-aware contingency planning. Partially observable Markov decision processes (POMDPs) and multistage stochastic programming formalize the principle that actions may depend only on information revealed so far (Kaelbling et al., 1998; Shapiro et al., 2021). Reachability and control barrier functions provide semantics for invariant or recoverable sets (Bansal et al., 2017; Bansal & Tomlin, 2021; Lin & Bansal, 2023; Ames et al., 2019; So &

Fan, 2023), while predictive/backup filters enforce online feasibility (Koller et al., 2018; Wabersich & Zeilinger, 2021; Tearle et al., 2021). Representative risk-aware comparators include robust scenario MPC (Batkovic et al., 2020), the Predictive Safety Filter (Wabersich & Zeilinger, 2021), the distributionally robust conditional value-at-risk (DR-CVaR) Safety Filter (Safaoui & Summers, 2024), and the Conformal Predictive Safety Filter (Strawn et al., 2023); more broadly, conditional value-at-risk (CVaR) and distributionally robust formulations control uncertain tails (Uryasev, 2000; Hakobyan & Yang, 2021; Nair et al., 2024). Multipolicy and Risk-aware Contingency planning (MARC) and Risk-Aware Contingency Planning (RACP) retain multiple policies under multimodal futures (Li et al., 2023; Mustafa et al., 2024). OC-RAP combines these themes but targets a specific failure mode: a safe option may exist in every hidden branch while no single option is valid across branches that are still observationally aliased. Its defining operation is therefore *aggregation before option selection*, not merely a more conservative scalar risk score.

Learning-based planning and closed-loop evaluation. Recent planners use interaction modeling, imitation, topology, diffusion, flow matching, or rule-aligned trajectory generation, including GameFormer (Huang et al., 2023), PlanTF (Cheng et al., 2024b), PLUTO (Cheng et al., 2024a), BeTopNet (Liu et al., 2024), Diffusion Planner (Zheng et al., 2025), Flow Planner (Tan et al., 2026), and Plan-R1 (Tang et al., 2026). Predictive Driver Model (PDM)-Closed/PDM-Hybrid (Dauner et al., 2023) and the Intelligent Driver Model (IDM) (Treiber et al., 2000) provide established non-learned references. nuPlan, CommonRoad, and Waymax support standardized or scalable closed-loop motion-planning evaluation (Caesar et al., 2021; Althoff et al., 2017; Gulino et al., 2023). These methods are strong behavioral planners or safety references, but they are not designed to certify that a recovery option is common to observationally indistinguishable latent futures. We therefore use representative planners, safety filters, contingency methods, and post-impact controllers as adapted comparators under a common executable interface rather than treating their native benchmark scores as directly transferable.

3 Problem Formulation

At planning time t, the planner observes the available history h_t and receives an exact nominal control prefix a_0 together with candidate prefixes $\mathcal{A}_t = \{a_0, a_1, \dots, a_N\}$. For a candidate a, the recovery model represents K plausible latent roots z_k with normalized probabilities p_k ($\sum_k p_k = 1$) and a finite recovery library $\mathcal{G} = \{g_1, \dots, g_L\}$. For each candidate–root–option triple (a, z_k, g_ℓ), the model assigns a signed recovery margin $M_{k\ell}(a)$: positive values denote recovery reserve, zero is the recovery boundary, and negative values denote recovery debt.

A hidden-root oracle may choose an option map $\pi_{\mathrm{orc}} : k \mapsto \ell$ using latent-root identity. A deployable map π_{dep} is restricted by the post-prefix observation and must therefore assign the same option to roots that remain observationally indistinguishable. Let $\Gamma_{\mathrm{dep}} \subseteq \Gamma_{\mathrm{orc}}$ denote these two policy classes, and let $p = (p_1, \dots, p_K)$. For any monotone lower-tail functional ϱ,

$$V_\varrho(a, \pi) = \varrho\big(\{M_{k,\pi(k)}(a)\}_{k=1}^K; p\big)\,, \tag{1}$$

$$V^\star_{\mathrm{dep}}(a) = \max_{\pi \in \Gamma_{\mathrm{dep}}} V_\varrho(a, \pi), \qquad V^\star_{\mathrm{orc}}(a) = \max_{\pi \in \Gamma_{\mathrm{orc}}} V_\varrho(a, \pi). \tag{2}$$

Proposition 1 (Information-pattern oracle dominance) *For every candidate a, $V^\star_{\mathrm{dep}}(a) \le V^\star_{\mathrm{orc}}(a)$; the inequality can be strict even if every latent root admits a positive recovery option.*

This result follows from policy-class inclusion and becomes strict when observationally aliased roots require incompatible recovery choices. OC-RAP estimates the deployable signed quantity while never exposing latent-root identity to the runtime selector.

4 OC-RAP

4.1 Executable Recovery and Runtime Factorization

OC-RAP separates three questions: *what can be executed*, *what recovery is legal under the deployed information pattern*, and *whether recovery evidence is sufficient to justify intervention*. Figure 1

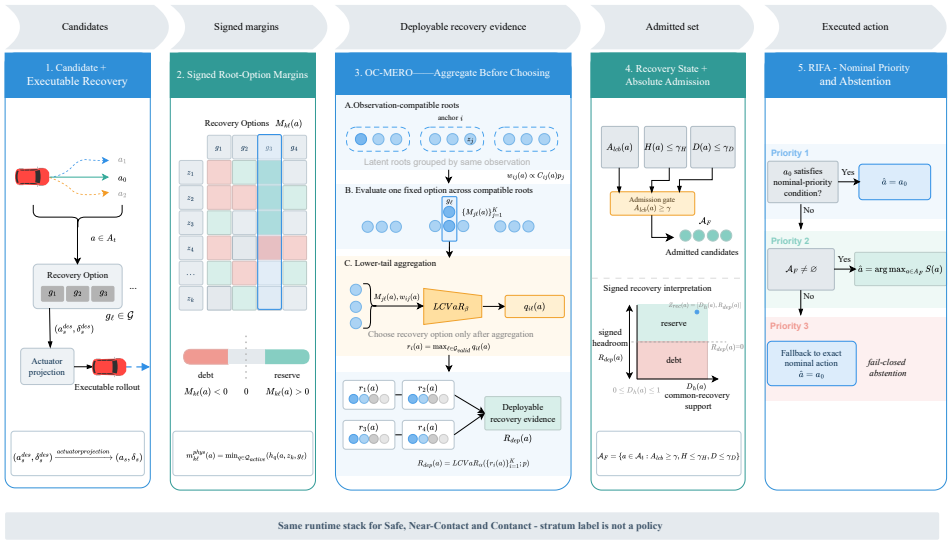


Figure 1: **OC-RAP runtime stack.** For each $a \in \mathcal{A}_t$, actuator-projected recovery options produce signed root–option margins $M_{k\ell}(a)$. OC-MERO aggregates each fixed option over observation-compatible roots to obtain $q_{i\ell}$, selects r_i only afterward, and then applies the outer aggregation to obtain $R_{\mathrm{dep}}(a)$. Conservative absolute evidence is combined with the candidate-generator validity flag $F_{\mathrm{pre}}(a)$, hard-violation score $H(a)$, and harm proxy $D(a)$ to define the admitted set; RIFA gives a_0 lexicographic priority, otherwise ranks only admitted candidates, and returns a_0 when none is admitted. The signed-recovery chart $Z_{\mathrm{rec}}(a) = [D_h(a), R_{\mathrm{dep}}(a)]$ is interpretive rather than an admission input. The learned recovery model and selector form are shared across Safe, Near-Contact, and Contact; only the pre-frozen scalar recovery-admission threshold is calibrated per evaluation stratum.

summarizes this dependency chain. Each recovery option is first projected into actuator-feasible acceleration, jerk, steering-angle, and steering-rate envelopes. The projected continuation is evaluated by normalized signed constraint functions $h_q(a, z_k, g_\ell)$ covering clearance, stopping, actuator control, route consistency, impact harm, stability, and secondary-collision risk; the deployed observation-only certificate additionally instantiates route and persistent-re-entry barriers when those constraints are active. In all cases $h_q \geq 0$ denotes satisfaction/reserve and $h_q < 0$ denotes violation/debt. Let $\mathcal{Q}_{\mathrm{active}}$ be the constraint indices active for the current recovery mode. We denote the direct non-compensatory physical bottleneck by

$$m^{\mathrm{phys}}_{k\ell}(a) = \min_{q \in \mathcal{Q}_{\mathrm{active}}} h_q(a, z_k, g_\ell), \tag{3}$$

so surplus on one physical dimension cannot compensate for a violation of another. The root–option margin $M_{k\ell}(a)$ used by OC-MERO is the recovery-model quantity supervised by the ordered structural contract in Appendix C.1; it is intentionally kept distinct from the direct physical bottleneck $m^{\mathrm{phys}}_{k\ell}(a)$. Projection details and candidate-relative constraint geometry are in Appendix B.

4.2 OC-MERO: Aggregate Before Choosing

Let $e^o_k(a) \in \mathbb{R}^{d_o}$ denote the predicted post-prefix observation embedding after candidate a under root z_k, where d_o is the embedding dimension. We use a soft compatibility kernel with temperature $\tau_o > 0$ and compatibility weights

$$C_{ij}(a) = \exp\left[-\frac{\|e^o_i(a) - e^o_j(a)\|_2^2/d_o}{\tau_o}\right], \qquad w_{ij}(a) = \frac{C_{ij}(a)p_j}{\sum_{u=1}^{K} C_{iu}(a)p_u + \varepsilon}, \tag{4}$$

where $\varepsilon > 0$ is a numerical stabilizer and $w_i = (w_{i1}, \ldots, w_{iK})$; the lower-tail routine normalizes w_i to unit mass before aggregation. Using lower-tail conditional value-at-risk (LCVaR) (Uryasev, 2000), with inner and outer tail masses $\beta, \alpha \in (0, 1]$, OC-MERO first evaluates every *fixed* option

over the compatible-root distribution for observation anchor i and only then chooses an option. Let $\mathcal{G}_{\text{valid}} \subseteq \mathcal{G}$ denote recovery options that pass the executable validity checks (candidate dependence is suppressed):

$$q_{i\ell}(a) = \mathrm{LCVaR}_{\beta}\big(\{M_{j\ell}(a)\}_{j=1}^{K}; w_i\big)\,, \tag{5}$$

$$r_i(a) = \max_{\ell \in \mathcal{G}_{\text{valid}}} q_{i\ell}(a), \qquad R_{\text{dep}}(a) = \mathrm{LCVaR}_{\alpha}\big(\{r_i(a)\}_{i=1}^{K}; p\big)\,. \tag{6}$$

Thus $q_{i\ell}$ is the conservative compatible-root value of option g_ℓ, r_i is the best such fixed-option value available at observation anchor i, and R_{dep} is the outer lower-tail deployable recovery value. This operator ordering enforces the information constraint: $q_{i\ell}$ is formed for a fixed option before the maximization over ℓ, so roots that remain indistinguishable to the deployed vehicle cannot authorize different recoveries.

For comparison, the implementation also reports a privileged *choose-first* oracle diagnostic: it selects the best valid option independently in each latent root and then evaluates the root-level lower tail,

$$R_{\text{orc}}(a) = \mathrm{LCVaR}_{\beta}\left(\left\{\max_{\ell \in \mathcal{G}_{\text{valid}}} M_{k\ell}(a)\right\}_{k=1}^{K}; p\right). \tag{7}$$

This diagnostic is distinct from the policy-class value $V^{\star}_{\text{orc}}$ in Eq. (2): because Eq. (7) and the nested deployable score in Eq. (6) use different aggregation structures, their raw difference is not itself an oracle upper-bound theorem. Runtime therefore uses only the one-sided discrepancy

$$G_{+}(a) = [R_{\text{orc}}(a) - R_{\text{dep}}(a)]_{+} \tag{8}$$

as a conservative diagnostic of choose-first optimism.

Theorem 1 (OC-MERO hard-partition measurability) *In the hard-partition limit, deterministic tie-breaking makes the option selected for each exact observation-equivalence class by Eq.* (6) *measurable with respect to the deployed post-prefix observation.*

The policy-class oracle dominance is Proposition 1; the measurability proof and the implementation-equivalent weighted lower-tail operator are in Appendices C and A.3.

4.3 Signed Recovery and RIFA

For each observation anchor, let $q_i^{\max}(a) \equiv r_i(a) = \max_{\ell \in \mathcal{G}_{\text{valid}}} q_{i\ell}(a)$ and define the supported mass $D_h(a) = \sum_i p_i \mathbf{1}[q_i^{\max}(a) \geq 0]$. We summarize recovery state by $Z_{\text{rec}}(a) = [D_h(a), R_{\text{dep}}(a)]$: D_h measures the probability mass of anchors whose compatible roots support a common fixed recovery option, while the sign of R_{dep} distinguishes reserve ($R_{\text{dep}} > 0$) from debt ($R_{\text{dep}} < 0$). This state chart is interpretive; RIFA admission below is determined by absolute evidence and candidate-prefix gates, not by Z_{rec} itself. Near-Contact therefore asks the planner to preserve or enlarge positive lower-tail headroom; Contact asks it to repay negative debt and achieve persistent re-entry. The continuous-time interpretation and additional state-chart quantities are deferred to Appendix B.

Recoverability alone does not establish that an intervention is preferable to nominal behavior. RIFA therefore constructs the conservative absolute-admission score

$$A_{\text{lcb}}(a) = \widehat{R}_{\text{dep}}(a) - \lambda_G \widehat{G}_{+}(a), \qquad \widehat{G}_{+}(a) = [\widehat{R}_{\text{orc}}(a) - \widehat{R}_{\text{dep}}(a)]_{+}, \tag{9}$$

where hats denote deployed recovery-model estimates rather than privileged truth-side values and the frozen selector uses $\lambda_G = 0.10$. For the current evaluation stratum b, let $\gamma_{\text{rec}}^{(b)}$ be its pre-frozen calibrated absolute-recovery threshold. The candidate generator additionally supplies a validity flag $F_{\text{pre}}(a) \in \{0, 1\}$, a nonnegative hard-violation score $H(a)$, and a nonnegative harm proxy $D(a)$; their explicit construction is given in Appendix B.4. RIFA admits only

$$\mathcal{A}_F^{(b)} = \{a \in \mathcal{A}_t : F_{\text{pre}}(a) = 1,\ A_{\text{lcb}}(a) \geq \gamma_{\text{rec}}^{(b)},\ H(a) \leq \gamma_H,\ D(a) \leq \gamma_D\}. \tag{10}$$

The frozen limits are $\gamma_H = 0$ and $\gamma_D = 5.0$. The exact nominal anchor has lexicographic priority when it passes the same prefix gates and either absolute admission or the fixed nominal slack rule; otherwise, the selector ranks only candidates in $\mathcal{A}_F^{(b)}$ using upstream utility, deviation cost, and recovery margin. If no candidate is admitted, strict RIFA conformance executes a_0. Hence

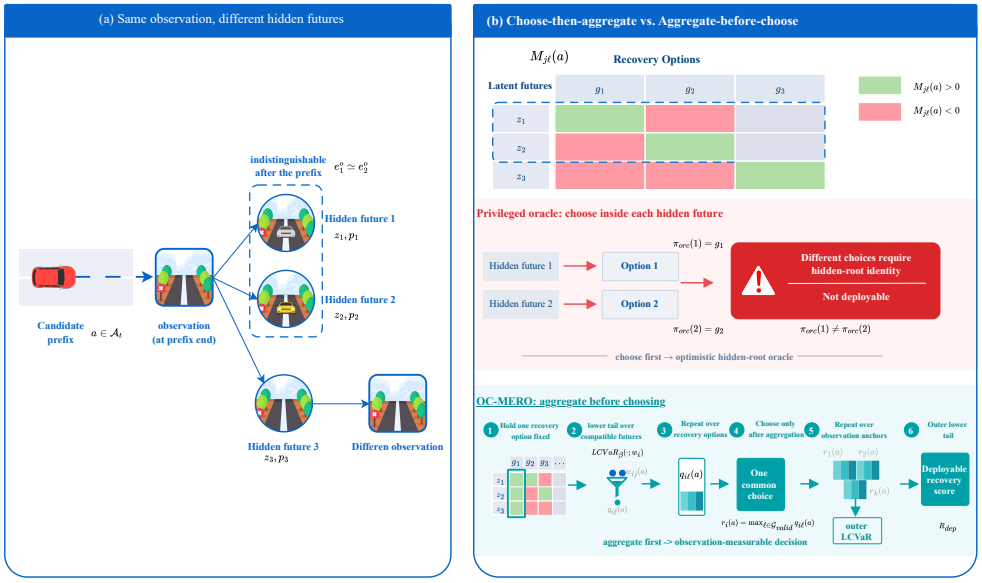


Figure 2: **Why operator order matters.** (a) Distinct latent roots can remain observation-compatible after the same prefix; $e_1^o(a) \simeq e_2^o(a)$ is schematic, while deployment uses the soft weights w_{ij}. (b) A choose-first oracle can use hidden-root identity to set $\pi_{\mathrm{orc}}(1) \neq \pi_{\mathrm{orc}}(2)$. OC-MERO instead holds g_ℓ fixed, forms $q_{i\ell}$ over compatible roots, then computes $r_i = \max_{\ell \in \mathcal{G}_{\mathrm{valid}}} q_{i\ell}$ and applies the outer lower-tail aggregation to obtain R_{dep}; the recovery choice is therefore observation-measurable in the hard-partition limit.

downstream preference scoring cannot rescue an absolute rejection. The complete score, prefix diagnostics, nominal slack rule, abstention proposition, runtime pseudocode, teacher semantics, and offline identifiability checks are in Appendix B. Recovery evidence and the prefix-admission diagnostics themselves use no truth-side future labels; Appendix B.5 records a separate benchmark-utility caveat discovered in the supplied final evaluation code.

5 Experiments

5.1 Protocol

We construct recovery strata from the Waymo Open Motion Dataset (WOMD) (Ettinger et al., 2021) and execute all closed-loop rollouts in Waymax (Gulino et al., 2023). The Safe stratum tests nominal non-interference, the Near-Contact stratum concentrates on low-headroom and observation-aliasing cases, and the Contact stratum starts from simulator-observed contact states. The same learned recovery model, candidate generator, recovery library, and selector functional form are used in all three strata; no stratum identity enters the learned recovery representation. The scalar admission threshold $\gamma_{\mathrm{rec}}^{(b)}$ is calibrated and frozen separately for each evaluation stratum before the final target-locked runs. Planner *route construction* uses only WOMD self-driving-car (SDC) path-connectivity geometry and the current simulated pose/heading; it does not use the logged-future route flag or logged future trajectory to select a route. A distinct inherited benchmark-reference term in the upstream candidate utility is disclosed in Appendix B.5.

Before final evaluation, we freeze a method-independent observation-legal target lock, yielding 174 Safe and 250 Near-Contact targets. For Contact, 747 eligible targets are advanced with an exact-a_0 prelude; one scene-disjoint observed-contact anchor with at least 40 remaining evaluation steps is frozen per eligible scene, yielding 8 fully paired anchors for the physical post-impact summary. Checkpoints, calibration, locks, metric code, and evaluation source are fixed thereafter. The statistical unit is one scene-level rollout. Clearance is the signed oriented-box distance; time-to-collision (TTC) is a constant-relative-velocity swept-box proxy. Contact metrics are measured only after an

Table 1: Near-Contact comparison on the common 250-scene lock. Here p05 denotes the 5th percentile; clearance p05 is over non-collision scenes, whereas TTC p05 uses the full cohort.

Method	Coll.%↓	Non-coll. clear. p05 (m)↑	TTC p05 (s)↑	Crit.-TTC exp. (s)↓	TTC-def. AUC (s^2)↓
OC-RAP Balanced	6.10	**0.661**	**0.130**	0.549	**0.630**
OC-RAP Precision	6.10	**0.661**	**0.130**	0.549	**0.630**
MARC	6.00	0.593	0.000	0.576	0.748
RACP	7.60	0.610	0.000	0.695	0.972
Robust scenario MPC	6.40	0.577	0.000	0.618	0.710
Predictive Safety Filter	18.80	0.576	0.000	0.921	1.488
DR-CVaR Safety Filter	6.40	0.600	0.000	**0.532**	0.725
Conformal Predictive Safety Filter	10.80	0.484	0.000	0.714	1.026
Flow Planner	**5.20**	0.620	0.022	0.564	0.682
Plan-R1	10.80	0.650	0.000	0.570	0.903
BeTopNet	10.80	0.600	0.000	0.595	0.933

observed-contact boundary, so collision/overlap at that boundary serves as a conditioning event rather than a pre-contact avoidance result. We report continuous re-contact diagnostics separately because their denominators are defined by event eligibility: the initial-separation rate (ISR) is defined over observed-contact rollouts, whereas the re-contact-after-separation rate (RASR), mean re-contact episode count (MREC), mean re-contact duration (MRD), and mean exposure ratio (MER) condition on successful initial separation. Full construction and metric definitions are in Appendices D and F.

We compare with learned planners, contingency/risk-aware methods, safety filters, and post-impact controllers cited in Section 2. Because several baselines were developed for nuPlan (Caesar et al., 2021), CommonRoad (Althoff et al., 2017), or higher-fidelity post-impact models, our rows are *principle-preserving adaptations* to a common candidate interface and Waymax execution path rather than claims to reproduce each method's native leaderboard result. Within a stratum, every method receives the same target lock, route contract, horizon, and metric implementation. Deployment latency includes state-history construction, candidate-feature computation, and policy selection, but excludes simulation, labels, metrics, rendering, and training. Baseline details and the complete Safe table are in Appendix G.

5.2 Closed-Loop Comparisons

We report both frozen OC-RAP operating points without post-hoc variant selection. We keep the three strata separate because they measure different questions: nominal preservation in Safe, lower-tail headroom in Near-Contact, and post-impact separation in Contact.

Safe. Appendix Table 6 shows that OC-RAP Balanced and Precision are the only rows with zero intervention and exact NUP $= 1$. Diffusion Planner has lower collision/off-road rates and PDM-Closed has lower latency, so the Safe results support a non-interference claim, not metric-wise dominance.

Near-Contact. OC-RAP obtains the largest non-collision clearance p05 (0.661 m), the highest TTC p05 (0.130 s), and the lowest TTC-deficit AUC (0.630 s^2). These gains do not constitute uniform dominance across safety metrics: Flow Planner has the lowest collision rate (5.20% versus 6.10%), and DR-CVaR has the shortest critical-TTC exposure (0.532 s versus 0.549 s). The results therefore favor OC-RAP in terms of lower-tail geometric headroom and integrated TTC deficit at this operating point, while leaving a small collision-rate gap to the best comparator.

Contact. On the small but fully paired 8-anchor physical cohort, OC-RAP has the shortest overlap (1.563 s), lowest penetration AUC (2.280 m s), and largest terminal clearance (19.007 m). Relative to the strongest non-OC-RAP result in each corresponding column, these values represent 5.3% shorter overlap, 9.7% lower penetration AUC, and 2.2% higher terminal clearance; PIB remains substantially faster (121.38 ms). Table 3 resolves a limitation of a single binary re-contact count by separating *whether recovery first separates* from *what happens after separation*. OC-RAP achieves an ISR of 88.5%, compared with the highest baseline ISR of 84.6%; among separation-eligible rollouts, it has the lowest RASR (25.0%, versus 31.8%), MRD (0.195 s, versus 0.218 s), and MER (8.6%, versus 9.4%). Its MREC advantage is smaller (0.350 versus 0.364) but is consistent in direction. Thus the

Table 2: Physical post-impact comparison on the common 8-scene observed-contact anchor lock. Lower overlap, penetration area under the curve (AUC), and latency are better; higher terminal clearance is better.

Method	Overlap dur. (s)↓	Penetr. AUC (m s)↓	Term. clear. (m)↑	Latency (ms)↓
OC-RAP Balanced	**1.563**	**2.280**	**19.007**	227.81
OC-RAP Precision	**1.563**	**2.280**	**19.007**	219.15
Integrated post-impact MPC	1.825	2.799	12.725	322.17
PIB	1.863	2.919	12.786	**121.38**
Post-impact motion planning + TVLQR	1.688	2.728	14.482	226.44
Post-collision trajectory restoration	1.650	2.526	18.596	126.73
FCC-MPC	1.738	2.706	17.794	132.66
Post-impact SMC + fault-tolerant QP	1.688	2.649	17.271	139.66

Table 3: Continuous re-contact diagnostics in Contact. ISR is the initial-separation rate over observed-contact rollouts. RASR, MREC, MRD, and MER are computed over separation-eligible rollouts; zero-re-contact rollouts contribute zero to MREC/MRD/MER. Higher ISR is better; all other columns are lower-is-better.

Method	ISR (%)↑	RASR (%)↓	MREC↓	MRD (s)↓	MER (%)↓
OC-RAP	**88.5**	**25.0**	**0.350**	**0.195**	**8.6**
Integrated post-impact MPC	80.8	36.4	0.378	0.218	10.7
PIB	76.9	31.8	0.364	0.235	9.4
Post-impact motion planning + TVLQR	84.6	33.3	0.381	0.241	12.6
Post-collision trajectory restoration	84.6	38.1	0.364	0.241	11.9
FCC-MPC	80.8	39.1	0.381	0.262	12.9
Post-impact SMC + fault-tolerant QP	80.8	42.9	0.418	0.290	14.1

Contact evidence supports both faster physical debt repayment and cleaner post-separation recovery, while the $n = 8$ physical endpoints should not be generalized to pre-contact collision frequency or population-level rates.

5.3 Matched Mechanism Ablations

Table 4 uses one-factor inference-time knockouts with the same frozen model, calibration, target lock, candidate budget, and evaluation semantics as the matched Full reference; no knockout is retrained or recalibrated. We interpret these results strictly within the ablation table and use Tables 1–3 for cross-method claims.

Removing observation consistency or replacing the nested lower tail with a mean degrades both physical and preservation endpoints: Near-Contact collision rises from 6.0% to 10.20% and 9.80%, respectively; TTC-deficit AUC rises from 0.6298 to 0.9544 and 0.9191 s^2, respectively; and intervention rises from 0.17% to 5.90% and 10.49%, respectively. Removing actuator projection or route alignment also raises collision to 10.60% and TTC-deficit AUC to 1.0067 s^2; in Contact, removing actuator projection, route alignment, or persistent re-entry reduces terminal clearance by roughly 1.1 m relative to the matched Full row and increases overlap and penetration.

RIFA is equally consequential for intervention control. Without absolute admission, Safe intervention rises to 9.37% and Near-Contact collision rises to 9.80%; without exact-nominal abstention, Safe intervention reaches 47.53% with NUP 0.8005, and Near-Contact intervention reaches 69.77% with NUP 0.6304. The latter knockout gives the largest Contact degradation in the table (14.160 m terminal clearance, 1.7875 s overlap, 2.7839 m s penetration AUC). These matched changes support the intended factorization: OC-MERO governs which recovery evidence is deployably credible, executable constraints govern what can physically recover, and RIFA prevents unsupported evidence from driving control interventions. Static target-locked rollouts are provided in Appendix G, and full-horizon synchronized videos are provided in the supplementary archive; both are qualitative only and are not used for quantitative claims.

Table 4: Matched frozen-module ablations. Lower collision/intervention/TTC-deficit/overlap/penetration is better; higher NUP/terminal clearance is better. "–" means the mechanism is inactive in that stratum.

	Safe		Near-Contact				Contact		
Configuration	Interv.%↓	NUP↑	Coll.%↓	Interv.%↓	NUP↑	TTC-def. AUC↓	Term. clear. (m)↑	Overlap (s)↓	Penetr. AUC (m s)↓
Full OC-RAP	0.00	1.0000	6.0	0.17	0.9993	0.6298	20.107	1.5014	2.1394
No observation consistency	–	–	10.20	5.90	0.9875	0.9544	19.052	1.5625	2.2777
Mean instead of lower tail	–	–	9.80	10.49	0.9582	0.9191	18.966	1.5875	2.3092
No actuator projection	–	–	10.60	0.12	0.9994	1.0067	19.007	1.5625	2.2796
No route alignment	–	–	10.60	0.28	0.9985	1.0067	19.007	1.5625	2.2796
No persistent re-entry	–	–	–	–	–	–	19.007	1.5625	2.2796
No RIFA absolute admission	9.37	0.9744	9.80	5.98	0.9708	0.9577	18.476	1.6625	2.5583
No exact-nominal abstention	47.53	0.8005	10.60	69.77	0.6304	1.0067	14.160	1.7875	2.7839

6 LIMITATIONS

OC-RAP can certify only recoveries represented by its finite libraries of candidates, latent roots, and recovery options, and its deployability statement is conditional on the learned observation-compatibility kernel and signed constraints. The hard observation partition gives an exact measurability statement; the deployed soft kernel is a relaxation. The finite-horizon recovery library is also not an infinite-horizon safety certificate. In addition, the final operating threshold is calibrated separately within each stratum, so we do not claim a universal deployment threshold.

The empirical evidence is deliberately narrower than a claim of overall planning superiority. Diffusion Planner has lower Safe collision/off-road rates; Flow Planner has the lowest Near-Contact collision rate; DR-CVaR has the shortest critical-TTC exposure; and the physical Contact comparison contains only eight fully paired anchors. The continuous re-contact statistics use event-eligible denominators and therefore should not be interpreted as simple fractions of the eight physical anchors. Moreover, a recoverable intervention need not be causally better than nominal behavior. RIFA handles this conservatively through nominal priority and abstention, but learning an identifiable treatment-benefit signal remains open.

7 CONCLUSION

OC-RAP frames recovery planning as an information-constrained decision problem. OC-MERO requires a recovery option to survive the lower tail of futures that remain observationally compatible before that option can be selected; actuator projection makes the certified continuation executable; the signed margin unifies reserve preservation before contact with debt repayment after contact; and RIFA keeps recovery admission separate from intervention preference. The final closed-loop results support this factorization while also exposing its trade-offs: exact Safe nominal preservation, stronger Near-Contact lower-tail headroom without collision-rate dominance, and improved post-impact separation and post-separation re-contact metrics on a contact-conditioned protocol.

AI USE STATEMENT

Generative AI tools were used to improve manuscript readability and organization, and to assist with the critique and presentation of experimental results. The proposed method, experimental setup, and reported numerical results were developed or obtained by the authors. All AI-assisted text was manually reviewed by the authors, and the technical claims, numerical results, and references were checked against the underlying experiments and sources. The authors take full responsibility for the final content of this work.

REPRODUCIBILITY STATEMENT

Sections 4 and Appendices A–F specify the operators, dataset construction, frozen evaluation protocol, and reporting semantics. The accompanying artifact contains the data, training, calibration, and evaluation interfaces used for the reported experiments.

REFERENCES

Matthias Althoff, Markus Koschi, and Stefanie Manzinger. Commonroad: Composable benchmarks for motion planning on roads. In *2017 IEEE intelligent vehicles symposium (IV)*, pp. 719–726. IEEE, 2017.

Aaron D Ames, Samuel Coogan, Magnus Egerstedt, Gennaro Notomista, Koushil Sreenath, and Paulo Tabuada. Control barrier functions: Theory and applications. In *2019 18th European control conference (ECC)*, pp. 3420–3431. Ieee, 2019.

Di Ao, Jialin Li, Lin Zhang, and Jinlong Hong. Advanced post-impact safety and stability control for electric vehicles. *IET Intelligent Transport Systems*, 16(12):1753–1767, 2022.

Somil Bansal and Claire J Tomlin. Deepreach: A deep learning approach to high-dimensional reachability. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1817–1824. IEEE, 2021.

Somil Bansal, Mo Chen, Sylvia Herbert, and Claire J Tomlin. Hamilton-jacobi reachability: A brief overview and recent advances. In *2017 IEEE 56th annual conference on decision and control (CDC)*, pp. 2242–2253. IEEE, 2017.

Ivo Batkovic, Ugo Rosolia, Mario Zanon, and Paolo Falcone. A robust scenario mpc approach for uncertain multi-modal obstacles. *IEEE Control Systems Letters*, 5(3):947–952, 2020.

Leonardo Bosia, Luca Manganotto, Marco Anghileri, Stefano Dolci, Paolo Astori, and Cing-Dao Kan. Real-time collision mitigation strategies for autonomous vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 25(10):13483–13493, 2024.

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.

Mingcong Cao, Chuan Hu, Rongrong Wang, Jinxiang Wang, and Nan Chen. Compensatory model predictive control for post-impact trajectory tracking via active front steering and differential torque vectoring. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 235(4):903–919, 2021.

Jie Cheng, Yingbing Chen, and Qifeng Chen. Pluto: Pushing the limit of imitation learning-based planning for autonomous driving. *arXiv preprint arXiv:2404.14327*, 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 *Conference on Robot Learning*, pp. 1268–1281. PMLR, 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.

Samsaptak Ghosh, M Felix Orlando, and Sohom Chakrabarty. Post-collision trajectory restoration for a single-track ackermann vehicle using heuristic steering and tractive force functions. *arXiv preprint arXiv:2602.08444*, 2026.

Cole Gulino, Justin Fu, Wenjie Luo, George Tucker, Eli Bronstein, Yiren Lu, Jean Harb, Xinlei Pan, Yan Wang, Xiangyu Chen, et al. Waymax: An accelerated, data-driven simulator for large-scale autonomous driving research. *Advances in Neural Information Processing Systems*, 36:7730–7742, 2023.

Astghik Hakobyan and Insoon Yang. Wasserstein distributionally robust motion control for collision avoidance using conditional value-at-risk. *IEEE Transactions on Robotics*, 38(2):939–957, 2021.

Zhiyu Huang, Haochen Liu, and Chen Lv. Gameformer: Game-theoretic modeling and learning of transformer-based interactive prediction and planning for autonomous driving. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 3880–3890. IEEE, 2023.

Leslie Pack Kaelbling, Michael L Littman, and Anthony R Cassandra. Planning and acting in partially observable stochastic domains. *Artificial intelligence*, 101(1-2):99–134, 1998.

Torsten Koller, Felix Berkenkamp, Matteo Turchetta, and Andreas Krause. Learning-based model predictive control for safe exploration. In *2018 IEEE conference on decision and control (CDC)*, pp. 6059–6066. IEEE, 2018.

Philip Koopman. Anatomy of a robotaxi crash: Lessons from the cruise pedestrian dragging mishap. In *International Conference on Computer Safety, Reliability, and Security*, pp. 119–133. Springer, 2024.

Daofei Li, Jiajie Zhang, Bin Xiao, Binbin Tang, and Zhaohan Hu. Vehicle crash mitigation strategy in unavoidable collision scenarios: Focusing on motion planning by considering a generalized crash severity model. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 44(12):581, 2022.

Tong Li, Lu Zhang, Sikang Liu, and Shaojie Shen. Marc: Multipolicy and risk-aware contingency planning for autonomous driving. *IEEE Robotics and Automation Letters*, 8(10):6587–6594, 2023.

Albert Lin and Somil Bansal. Generating formal safety assurances for high-dimensional reachability. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 10525–10531. IEEE, 2023.

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.

Jianbo Lu, Hassen Hammoud, Todd Clark, Otto Hofmann, Mohsen Lakehal-ayat, Shweta Farmer, Jason Shomsky, and Roland Schaefer. A system for autonomous braking of a vehicle following collision. In *WCX™ 17: SAE World Congress Experience*. SAE Technical Paper, 2017.

Khaled A Mustafa, Daniel Jarne Ornia, Jens Kober, and Javier Alonso-Mora. Racp: Risk-aware contingency planning with multi-modal predictions. *IEEE Transactions on Intelligent Vehicles*, 10 (1):228–243, 2024.

Siddharth H Nair, Hotae Lee, Eunhyek Joa, Yan Wang, H Eric Tseng, and Francesco Borrelli. Predictive control for autonomous driving with uncertain, multimodal predictions. *IEEE transactions on control systems technology*, 33(4):1178–1192, 2024.

Masoumeh Parseh, Mikael Nybacka, and Fredrik Asplund. Motion planning for autonomous vehicles with the inclusion of post-impact motions for minimising collision risk. *Vehicle system dynamics*, 61(6):1707–1733, 2023.

Sleiman Safaoui and Tyler H Summers. Distributionally robust cvar-based safety filtering for motion planning in uncertain environments. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 103–109. IEEE, 2024.

Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczynski. *Lectures on stochastic programming: modeling and theory*. SIAM, 2021.

Oswin So and Chuchu Fan. Solving stabilize-avoid via epigraph form optimal control using deep reinforcement learning. In *Robotics: Science and Systems*, 2023.

Kegan J Strawn, Nora Ayanian, and Lars Lindemann. Conformal predictive safety filter for rl controllers in dynamic environments. *IEEE Robotics and Automation Letters*, 8(11):7833–7840, 2023.

Tianyi Tan, Yinan Zheng, Ruiming Liang, Zexu Wang, Kexin Zheng, Jinliang Zheng, Jianxiong Li, Xianyuan Zhan, and Jingjing Liu. Flow matching-based autonomous driving planning with advanced interactive behavior modeling. *Advances in Neural Information Processing Systems*, 38: 38310–38335, 2026.

Xiaolong Tang, Meina Kan, Shiguang Shan, and Xilin Chen. Plan-r1: Safe and feasible trajectory planning as language modeling. In *International Conference on Learning Representations*, volume 2026, pp. 23715–23735, 2026.

Ben Tearle, Kim P Wabersich, Andrea Carron, and Melanie N Zeilinger. A predictive safety filter for learning-based racing control. *IEEE Robotics and Automation Letters*, 6(4):7635–7642, 2021.

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

Stanislav Uryasev. Conditional value-at-risk: Optimization algorithms and applications. In *proceedings of the IEEE/IAFE/INFORMS 2000 conference on computational intelligence for financial engineering (CIFEr)(Cat. No. 00TH8520)*, pp. 49–57. IEEE, 2000.

Kim Peter Wabersich and Melanie N Zeilinger. A predictive safety filter for learning-based control of constrained nonlinear dynamical systems. *Automatica*, 129:109597, 2021.

Cong Wang, Zhenpo Wang, Lei Zhang, Huilong Yu, and Dongpu Cao. Post-impact motion planning and tracking control for autonomous vehicles. *Chinese Journal of Mechanical Engineering*, 35(1): 54, 2022.

Cong Wang, Zhenpo Wang, Zhiqiang Zhang, Jizheng Liu, Wenbo Li, Yang Wu, Xiaoyu Li, Huilong Yu, and Dongpu Cao. Integrated post-impact planning and active safety control for autonomous vehicles. *IEEE Transactions on Intelligent Vehicles*, 8(3):2062–2076, 2023.

Cong Wang, Zhenpo Wang, Lei Zhang, Jun Chen, and Dongpu Cao. Post-impact stability control for road vehicles: State-of-the-art methodologies and perspectives. *IEEE Transactions on Intelligent Transportation Systems*, 25(8):8295–8312, 2024.

Derong Yang, Tim Gordon, and Ulrich Sander. Post-impact motion control of road vehicles. *Vehicle System Dynamics*, 63(7):1300–1357, 2025.

Yinan Zheng, Ruiming Liang, Kexin Zheng, Jinliang Zheng, Liyuan Mao, Jianxiong Li, Weihao Gu, Rui Ai, Shengbo Li, Xianyuan Zhan, et al. Diffusion-based planning for autonomous driving with flexible guidance. In *International conference on learning representations*, volume 2025, pp. 37207–37227, 2025.

A Implementation Details

A.1 Actuator Projection

At recovery-rollout step s, an option produces desired longitudinal acceleration and steering $(a_s^{\mathrm{des}}, \delta_s^{\mathrm{des}})$; the scalar a_s here is distinct from the candidate prefix a. With time step Δt, acceleration bounds $[a_{\min}, a_{\max}]$, jerk bound $j_{\max}$, steering-angle bound $\delta_{\max}$, and steering-rate bound $\dot{\delta}_{\max}$, let $\mathrm{clip}_{[l,u]}(x) = \min(u, \max(l, x))$ denote scalar clipping to $[l, u]$. The executable commands are

$$a_s = \mathrm{clip}_{[\max(a_{\min}, a_{s-1} - j_{\max}\Delta t),\, \min(a_{\max}, a_{s-1} + j_{\max}\Delta t)]}(a_s^{\mathrm{des}}), \tag{11}$$

$$\delta_s = \mathrm{clip}_{[\max(-\delta_{\max}, \delta_{s-1} - \dot{\delta}_{\max}\Delta t),\, \min(\delta_{\max}, \delta_{s-1} + \dot{\delta}_{\max}\Delta t)]}(\delta_s^{\mathrm{des}}). \tag{12}$$

The rollout and all downstream recovery constraints are computed from the projected commands. The finite recovery library $\mathcal{G}$ instantiates parameterized variants of eight semantic modes: stop, brake-in-lane, lateral escape, yield-and-rejoin, pull over, mitigate contact, post-contact stabilization, and secondary-collision avoidance. Mode-specific availability checks define $\mathcal{G}_{\mathrm{valid}}$ for each candidate (for example, pull-over requires an available shoulder and lateral/secondary-avoidance modes require adjacent space). The closed-loop evaluator exposes up to 12 valid option instances per planning step.

A.2 Observation-Only Route and Persistent Re-entry Barriers

Two option-level coordinates make the executable certificate explicit. Let y_s be the recovery trajectory's route-normal coordinate, $d_{\text{route,max}}$ the allowed lateral deviation, and $s_{\text{route}} > 0$ its normalization scale. For modes in which route consistency is active,

$$h_{\text{route}} = \tanh\left(\frac{d_{\text{route,max}} - \max_s |y_s|}{s_{\text{route}}}\right), \tag{13}$$

and post-contact stabilization treats this coordinate as inactive. For persistent re-entry, let c_s be the minimum signed clearance to the current-observation constant-velocity occupancy forecast along the projected continuation and let $\rho_s = c_s - d_{\text{safe}}$ be the signed safe-clearance reserve. If the observed prefix is already in contact/overlap, or the recovery mode is post-contact stabilization or secondary-collision avoidance, define

$$s^\star = \min\{s : \rho_s \geq 0\}, \qquad \rho_{\text{pers}} = \begin{cases} \min_{u \geq s^\star} \rho_u, & s^\star \text{ exists}, \\ \rho_T, & \text{otherwise}, \end{cases} \tag{14}$$

with ρ_{pers} set to a satisfied sentinel when the check is inactive. The certificate coordinate is

$$h_{\text{re}} = \tanh(\rho_{\text{pers}}/s_d), \tag{15}$$

where $s_d > 0$ is the distance normalization. Thus initial penetration does not cause permanent rejection, but a trajectory that re-enters and later re-contacts acquires negative persistent reserve. These quantities are computed from the executable recovery trajectory and the same current-observation occupancy forecast as the clearance witness; no hidden intent or truth-side route-blocked flag is read.

A.3 Weighted Lower-Tail Operator

For scores s_i and nonnegative weights w_i, the implementation first normalizes w to unit mass, then uses a stable sort to order the scores from worst to best and integrates exactly an α fraction of probability mass, including a fractional portion of the final item when necessary. An equivalent variational form is

$$\text{LCVaR}_\alpha(s; w) = \max_{\nu \in \mathbb{R}} \left[\nu - \frac{1}{\alpha} \sum_i w_i [\nu - s_i]_+\right]. \tag{16}$$

The implementation uses $\alpha = \beta = 0.2$ and up to the eight most compatible roots when compatibility sparsification is enabled.

B Additional Method Details

B.1 Candidate-Relative Constraint Geometry

For root z_k and option g_ℓ, define the candidate-relative response

$$\Delta h_q(a; a_0, z_k, g_\ell) = h_q(a, z_k, g_\ell) - h_q(a_0, z_k, g_\ell). \tag{17}$$

It is zero for the nominal prefix and positive when the candidate moves a constraint toward reserve or away from debt. For a generic state/feature point x and perturbation direction d, let $m(x) = \min_q h_q(x)$ with active set $\mathcal{I}(x) = \{q : h_q(x) = m(x)\}$. The one-sided directional derivative is

$$Dm(x)[d] = \min_{q \in \mathcal{I}(x)} \nabla h_q(x)^\top d. \tag{18}$$

Thus the limiting constraint identifies the boundary, while the useful action direction depends on how a candidate changes the signed active-constraint geometry.

B.2 Recovery State, Reserve, and Debt

For $q_i^{\max}(a) \equiv r_i(a) = \max_{\ell \in \mathcal{G}_{\text{valid}}} q_{i\ell}(a)$, define

$$D_h(a) = \sum_i p_i \mathbf{1}[q_i^{\max}(a) \geq 0], \qquad D_s(a) = \sum_i p_i \,\text{sigmoid}(q_i^{\max}(a)/\tau_q), \tag{19}$$

$$P_{\text{dep}}(a) = \text{sigmoid}(R_{\text{dep}}(a)), \qquad Q_{\text{gap}}(a) = \exp(-G_+(a)), \tag{20}$$

where $\mathrm{sigmoid}(x) = 1/(1+e^{-x})$ and $\tau_q > 0$ controls the soft support approximation. Here D_h and D_s are hard and soft common-recovery support, respectively; P_{dep} is a bounded visualization of deployable recovery, and Q_{gap} decreases with the one-sided choose-first discrepancy G_+ from Eq. (8). The deployed state chart is $Z_{\mathrm{rec}}(a) = [D_h(a), R_{\mathrm{dep}}(a)]$. D_h summarizes the mass of anchors with an observation-consistent common recovery; conditional on support, $R_{\mathrm{dep}} > 0$ is reserve and $R_{\mathrm{dep}} < 0$ is debt. Candidate motion is $\Delta Z_{\mathrm{rec}}(a; a_0) = Z_{\mathrm{rec}}(a) - Z_{\mathrm{rec}}(a_0)$.

B.3 Identifiable Absolute Recoverability

The physical margin in Eq. (3) is kept distinct from structural teacher rules, including recovery-mode validity, route blocking, secondary-risk logic, and rules for deliberately mined hidden branches. Structural or future metadata used for truth-side audits is never a runtime feature. Candidate-relative root-local analyses are authorized only after establishing split/merge-compatible counterfactual correspondence within shared exogenous/observation-compatible realization classes (Appendix E). Finally, for any scalar shift c, lower-tail CVaR is translation equivariant,

$$\mathrm{LCVaR}_\alpha(s + c\mathbf{1}; w) = \mathrm{LCVaR}_\alpha(s; w) + c, \tag{21}$$

so an option-wise scalar shift cannot change weak-root ordering or lower-tail ownership. These checks motivate introducing response capacity only after the correspondence associated with the recovery boundary is identifiable.

B.4 RIFA Score and Runtime Rule

The frozen candidate generator supplies three prefix-level quantities in addition to the recovery-model outputs. Let $I_{\mathrm{coll}}(a)$ and $I_{\mathrm{contact}}(a)$ denote its distance-proxy collision and contact indicators (thresholds 1.8 m and 0.8 m), $I_{\mathrm{off}}(a)$ the off-road indicator ($\max_s |y_s| > 6$ m), and $I_{\mathrm{wrong}}(a)$ the wrong-way indicator ($\cos(\psi_s - \psi_{\mathrm{route,s}}) < -0.2$ at any prefix step). Then

$$H(a) = I_{\mathrm{coll}}(a) + 2I_{\mathrm{off}}(a) + 2I_{\mathrm{wrong}}(a), \qquad D(a) = I_{\mathrm{contact}}(a) + 0.2I_{\mathrm{coll}}(a). \tag{22}$$

The Boolean $F_{\mathrm{pre}}(a)$ is true only when the prefix states/controls are finite, the generator's control-tolerance check passes, route topology is valid, and neither off-road nor wrong-way hard failure occurs. These are candidate-prefix diagnostics, not learned recovery heads and not truth-side future labels.

For a prefix with planar states $x_s^a \in \mathbb{R}^2$, the selector uses the normalized nominal deviation

$$d(a, a_0) = \frac{1}{5}\sqrt{\frac{1}{T}\sum_{s=1}^{T} \|x_s^a - x_s^{a_0}\|_2^2}, \tag{23}$$

where T is the common compared prefix length. The within-admission score for stratum b is

$$S^{(b)}(a) = U(a) - \lambda_{\mathrm{int}}\mathbf{1}[a \neq a_0] - \lambda_{\mathrm{dev}} d(a, a_0) + \lambda_{\mathrm{rec}}(A_{\mathrm{lcb}}(a) - \gamma_{\mathrm{rec}}^{(b)}), \tag{24}$$

where $U(a)$ is upstream planner utility. The frozen values are $\lambda_G = 0.10$, $\lambda_{\mathrm{int}} = 0.03$, $\lambda_{\mathrm{dev}} = 0.15$, and $\lambda_{\mathrm{rec}} = 0.02$, with $\gamma_H = 0$ and $\gamma_D = 5.0$. The only bucket-specific selector scalar is $\gamma_{\mathrm{rec}}^{(b)}$, loaded from the frozen calibration file for Safe, Near-Contact, or Contact.

Let $\widehat{G}_+(a) = [\widehat{R}_{\mathrm{orc}}(a) - \widehat{R}_{\mathrm{dep}}(a)]_+$ as in Eq. (9). The nominal prefix has lexicographic priority whenever it passes F_{pre}, H, and D and either (i) satisfies $A_{\mathrm{lcb}}(a_0) \geq \gamma_{\mathrm{rec}}^{(b)}$, or (ii) satisfies the fixed slack condition

$$A_{\mathrm{lcb}}(a_0) - \gamma_{\mathrm{rec}}^{(b)} \geq -0.03, \qquad \widehat{G}_+(a_0) \leq 0.50. \tag{25}$$

Otherwise, the publication selector chooses the highest-scoring candidate in $\mathcal{A}_F^{(b)}$; if $\mathcal{A}_F^{(b)} = \varnothing$, strict absolute-admission conformance returns a_0. The optional learned relative-policy heads in the code path are disabled in the frozen publication selector, so the reported runs use Eq. (24) only after absolute admission. We denote the selector output by $\hat{a}$.

Proposition 2 (Admission and abstention invariance) *An absolutely rejected non-nominal candidate cannot be selected; a prefix-safe nominal action satisfying the nominal-priority condition is selected; and if $\mathcal{A}_F^{(b)} = \varnothing$, the executed action is a_0.*

Algorithm 1 Receding-horizon OC-RAP

Observe h_t; generate a_0 and candidates $\mathcal{A}_t$.
for each $a \in \mathcal{A}_t$ **do**
  Project recovery options; roll out signed constraints; compute OC-MERO and $A_{\mathrm{lcb}}(a)$; read $F_{\mathrm{pre}}(a), H(a), D(a)$ from the candidate prefix.
end for
Form $\mathcal{A}_F^{(b)}$ using Eq. (10).
if a_0 satisfies the nominal-priority condition **then**
  $\hat{a} \leftarrow a_0$.
else if $\mathcal{A}_F^{(b)} \neq \varnothing$ **then**
  $\hat{a} \leftarrow \arg\max_{a \in \mathcal{A}_F^{(b)}} S^{(b)}(a)$.
else
  $\hat{a} \leftarrow a_0$.
end if
Execute $\hat{a}$ and replan.

B.5 Benchmark-Utility Deployment Caveat

The final evaluation code inherits the candidate generator's utility

$$U(a) = \Delta x(a) - 0.05C(a) - 0.5d_{\mathrm{route}}(a) - 0.05d_{\log}(a) - 5I_{\mathrm{off}}(a) - 5I_{\mathrm{wrong}}(a), \tag{26}$$

where $C(a)$ is the acceleration/jerk comfort cost and $d_{\log}(a)$ is terminal planar distance to the benchmark logged ego future. This low-weight reference term is used only after absolute recovery admission, but the supplied Waymax history path populates the future side from the logged trajectory. Consequently, the current frozen closed-loop numbers should be interpreted as *benchmark characterization*, not as a fully observation-only deployment claim for the complete selector. Strict deployment requires setting the logged-future utility weight to zero (or replacing it with an observation-legal nominal-reference term) and re-running the final characterization. This caveat does not change OC-MERO, the recovery certificate, or the absolute-before-relative RIFA ordering; it identifies an inherited upstream utility input that must be removed before making the stronger deployment claim.

B.6 Signed Viability Interpretation

The signed-margin convention has a continuous-time viability interpretation: before violation, useful actions preserve or increase positive reserve; after violation, they produce a positive signed response that repays the controlling negative debt. For a differentiable limiting signed constraint $h(t)$, where $h > 0$ denotes reserve and $h < 0$ denotes debt:

Theorem 2 (Signed invariance and finite-time debt repayment) *If $h \geq 0$ implies $\dot{h} \geq -\lambda h$ for $\lambda > 0$, then any trajectory with $h(0) \geq 0$ remains nonnegative while the condition holds. If $h < 0$ and $\dot{h} \geq c[-h]^{\eta}$ for $c > 0$ and $0 \leq \eta < 1$, then $d(t) = [-h(t)]_+$ reaches zero no later than*

$$T_{\mathrm{re}} \leq \frac{d(0)^{1-\eta}}{c(1-\eta)}. \tag{27}$$

For reference, the two differential conditions are $\dot{h} \geq -\lambda h$ on the nonnegative side and $\dot{h} \geq c[-h]^{\eta}$ on the negative side; we label them Eqs. (28)–(29) below for the proof.

$$\dot{h}(t) \geq -\lambda h(t), \quad h(t) \geq 0, \tag{28}$$

$$\dot{h}(t) \geq c[-h(t)]^{\eta}, \quad h(t) < 0. \tag{29}$$

The finite recovery library uses actuator projection, signed rollout constraints, and persistent re-entry checks as finite-horizon proxies for this interpretation; it is not an infinite-horizon certificate.

C Proofs and Structural Properties

Proof of Proposition 1. By definition, every observation-measurable recovery map is also a root-indexed map, hence $\Gamma_{\mathrm{dep}} \subseteq \Gamma_{\mathrm{orc}}$. Maximizing the same value functional over a subset cannot exceed

the maximum over the superset. Strictness is possible when two observationally equivalent roots have disjoint maximizing option sets: the oracle can choose separately, whereas any deployable map must use one common option. □

Active-constraint directional derivative. For $m(x) = \min_q h_q(x)$, write $h_q(x + \epsilon d) = h_q(x) + \epsilon \nabla h_q(x)^\top d + o(\epsilon)$. Constraints outside $\mathcal{I}(x)$ have a positive slack gap and cannot become minimizing for sufficiently small positive ϵ. Taking the minimum over the active first-order expansions gives Eq. (18).

Proof of Theorem 1. In the hard-partition limit, two anchors in the same exact post-prefix observation class have the same compatible-root set and the same normalized compatible-root weights. Therefore every fixed-option value $q_{i\ell}$ is identical for anchors in that class. The vector $(q_{i1}, \ldots, q_{iL})$ is consequently a function of the observation class alone; deterministic tie-breaking makes $\arg\max_{\ell \in \mathcal{G}_{\text{valid}}} q_{i\ell}$ a function of that class rather than hidden root identity. □

Proof of Proposition 2. The selector evaluates non-nominal preference only on $\mathcal{A}_F^{(b)}$, so a rejected non-nominal candidate cannot be returned. The nominal prefix-gate and admission/slack check is evaluated before non-nominal scoring and therefore has lexicographic priority. If $\mathcal{A}_F^{(b)}$ is empty, the strict-conformance branch executes a_0. □

Translation equivariance and weak-tail membership. Using the variational form in Eq. (16), substitute $s_i + c$ and change variables $\nu' = \nu - c$ to obtain $\mathrm{LCVaR}_\alpha(s + c\mathbf{1}; w) = c + \mathrm{LCVaR}_\alpha(s; w)$. Adding the same scalar preserves all pairwise differences $s_i - s_j$, hence the stable root order and fractional lower-tail membership are unchanged.

Proof of Theorem 2. When $h \geq 0$, Eq. (28) implies $\frac{d}{dt}(e^{\lambda t} h(t)) \geq 0$, hence $h(t) \geq h(0) e^{-\lambda t} \geq 0$. When $h < 0$, let $d = -h > 0$. Equation (29) gives $\dot{d} \leq -c d^\eta$. For $0 \leq \eta < 1$, integration yields $d(t)^{1-\eta} \leq d(0)^{1-\eta} - c(1-\eta)t$, so d reaches zero no later than Eq. (27). Applying the nonnegative-side condition thereafter prevents re-entry from being immediately lost. □

Mass conservation of the split/merge transport. For the equivalence-class transport in Appendix E,

$$\sum_j \Pi_e(k, j) = \min(c_e, n_e) \frac{c_{ke}}{c_e}, \qquad \sum_k \Pi_e(k, j) = \min(c_e, n_e) \frac{n_{je}}{n_e}. \tag{30}$$

Thus all common class mass is transported and the construction is invariant to permutations, splits, or merges of root slots inside the same exogenous/observation-compatible class. When $c_e = n_e$, the candidate and nominal class marginals are exactly preserved.

C.1 Physical and Structural Teacher Semantics

The physical margin before structural rules is Eq. (3). The stored teacher target is produced by an ordered structural contract rather than by an unconditional minimum of active slacks. In the supplied teacher, non-hidden post-contact stabilization, yield-and-rejoin, and pull-over modes are floored at 0.6; a route-blocked yield-and-rejoin is capped at -0.8; a secondary-threat/secondary-avoidance pair is floored at 0.9; deliberately mined hidden branches may use branch-specific replacement; and an invalid option receives a large negative sentinel. These operations encode structural deployability/admissibility semantics rather than additional physical clearance.

For absolute-source supervision, structurally mixed targets constrain the underlying physical signed margin rather than being treated as point-identified physical values. Teacher-positive recovery evidence may therefore contain physical recovery information without being interpreted as exact physical truth. Interval-valued truth-side audits derived from the ordered structural contract are used only for supervision and identifiability checks, not as privileged runtime inputs. The distinction does not depend on a particular truth-side parameterization, and structural or future metadata used to construct these audits never enters the deployed model.

D Dataset Construction and Audit

D.1 Source Data and Runtime Backend

The constructor reads WOMD motion v1.3.1 records and uses the Waymax closed-loop simulation backend (Ettinger et al., 2021; Gulino et al., 2023). Training buckets are sourced from WOMD training records; validation, calibration, and Safe/Near-Contact/Contact evaluation buckets are scene-separated slices constructed from the standard WOMD validation records. The official WOMD test split is not used for this construction, so we describe the benchmark as *WOMD-derived with Waymax execution*, not as an official WOMD test benchmark. Final reported runs use immutable checkpoints and calibration together with a method-independent observation-legal target lock; route-ineligible targets are excluded before method comparison rather than repaired with future labels.

Each scene–time group starts from a budget of 24 candidate proposals, 8 latent roots, and 12 recovery options. Dataset-quality filtering then retains at most 8 candidate prefixes for Safe and Near-Contact and at most 9 for Contact; the online closed-loop evaluator may still score the full 24-candidate runtime budget. The audit reports confirm Waymax runtime coverage for all audited futures. Safe buckets contain one replay future plus two reactive futures. Near-Contact uses these replay and reactive futures plus targeted hidden-yield, hidden-acceleration, low-friction-braking, and control-delay hypotheses. Contact additionally includes contact-impulse and secondary-collision-approach hypotheses. The constructor can also add visible perturbation roots where specified by the regime recipe.

D.2 Regime-Specific Construction

Safe. Safe emphasizes nominal behavior and non-interference. The training and development-adjudication recipes disable deliberately targeted futures and oracle-artifact mining, preserve a nominal candidate at each scene-time group, and exclude Near-Contact, Post-Contact, and oracle-artifact nominal samples under the paper regime definitions.

Near-Contact. Near-Contact uses interaction-biased time sampling, eight targeted futures, augmented hidden roots, visible perturbation roots, and balanced artifact and non-artifact sampling. Targeted hidden futures include vehicle yield/acceleration, low-friction braking, and control-delay noise. Artifact-pair construction is used to create cases in which branch-wise recovery can differ from deployable observation-consistent recovery.

Contact. Contact uses denser interaction-biased sampling and ten targeted futures. In addition to the Near-Contact hypotheses, it includes contact-impulse and secondary-collision-approach surrogates. The construction retains nominal candidates while emphasizing post-contact, oracle-artifact, re-contact, and secondary-risk cases.

D.3 Dataset Roles and Audited Statistics

The data are organized into train, validation, calibration, and development-adjudication roles for each regime. Calibration is held out from training and used only to determine operating points. The twelve canonical buckets pass schema and finite-value checks; stored and recomputed OC-MERO values agree to numerical precision, and per-role leakage checks report no leakage. Cross-role scene separation is enforced by the released split manifests. Table 5 summarizes these buckets.

The Safe buckets intentionally contain no oracle artifacts, yet they are not trivial all-positive sets: examples with negative deployable recoverability remain at approximately 5–10%. Near-Contact and Contact contain substantially more negative deployability and oracle artifacts, creating useful separation between branch-wise recoverability and observation-consistent deployability. Across the audited buckets, the average number of valid roots is approximately two in Safe, 6.4–6.5 in Near-Contact, and eight in Contact, matching the intended increase in latent ambiguity and post-contact structure. The 175-scene audited Safe test bucket in Table 5 precedes the final observation-legal route lock; one route-ineligible target is excluded method-independently, giving the 174-scene Safe cohort used in Table 6.

Table 5: Audited OC-RAP dataset buckets. "Neg. dep." is the fraction with negative deployable recoverability; "Artifact" is the oracle-artifact fraction; "Alias conflict" is the fraction of observation-compatible root pairs with incompatible best recovery options.

Role	Regime	Samples	Scenes	Groups	Neg. dep. (%)	Artifact (%)	Oracle rec. (%)	Alias conflict (%)
Train	Safe	20,000	1,171	2,500	9.88	0.00	90.11	0.00
Train	Near	13,324	600	1,800	55.31	18.94	63.63	16.15
Train	Contact	16,790	500	2,000	54.35	16.62	62.27	9.48
Val.	Safe	2,328	132	291	7.35	0.00	92.65	0.00
Val.	Near	3,445	176	433	50.42	24.59	74.17	20.39
Val.	Contact	6,477	211	723	46.15	21.86	75.71	13.54
Calib.	Safe	2,544	135	318	5.35	0.00	94.65	0.00
Calib.	Near	6,039	316	765	44.83	23.99	79.17	19.57
Calib.	Contact	16,843	543	1,896	41.69	21.21	79.52	14.23
Test	Safe	3,216	175	402	6.90	0.00	93.10	0.00
Test	Near	4,723	250	595	48.80	24.41	75.61	20.92
Test	Contact	6,687	209	747	44.40	21.80	77.40	14.05

D.4 Reproducibility Notes

The reproducibility package separates train, calibration, and validation/evaluation construction and records the source WOMD files, scenario budgets, candidate/root/option counts, time-sampling quotas, targeted-future types, artifact mining rules, and regime filters. The accompanying audit reports and split manifests make role partitioning, scene separation, and operating-point provenance independently auditable.

E Offline Identifiability Protocol

The runtime planner does not consume latent future identity. Offline mechanism analysis follows the same principle. When comparing a candidate with the nominal action, latent roots are not matched by an arbitrary root-slot index. Instead, future instances are grouped by observation-compatible branch/exogenous realization semantics, and probability mass is transported only within a shared equivalence class. For an equivalence class e present on both sides, let c_{ke} and n_{je} be candidate and nominal root masses and let $c_e = \sum_k c_{ke} > 0$, $n_e = \sum_j n_{je} > 0$. A split/merge-compatible transport is

$$\Pi_e(k, j) = \min(c_e, n_e) \frac{c_{ke}}{c_e} \frac{n_{je}}{n_e}, \qquad \Pi = \sum_e \Pi_e. \tag{31}$$

Missing exogenous fingerprints are treated as unresolved rather than silently matched. This transport is used for identifiability and audit analyses only; it is not a runtime input or an additional learned planner module. The purpose is methodological: to establish a deployable correspondence before fitting higher-capacity response models, rather than allowing hidden future identity to leak into the learned recovery source.

F Additional Reporting Details

Metric semantics. Signed clearance is the exact minimum distance between oriented ego and agent boxes, and is negative under penetration. Time-to-collision (TTC) is computed with a swept separating-axis test under constant relative velocity and frozen footprint/heading over the local prediction interval; it is therefore a geometric TTC proxy rather than an official Waymax metric. Exposure durations and deficit areas under the curve (AUCs) use the simulator sampling interval without adding a fictitious terminal interval. Route progression projects the simulated start and end self-driving-car (SDC) states onto the fixed observation-legal route; it is distinct from any logged-future progression label. Bounded nominal-utility preservation (NUP) measures preservation of the nominal utility anchor and equals one by construction for the exact nominal control, so we report it as an intervention/preservation diagnostic rather than as a standalone safety guarantee.

For continuous re-contact evaluation, let $\mathcal{S}$ be the observed-contact rollouts and $\mathcal{S}_{\text{sep}}$ the subset that enters `overlap = false` at least once after the initial continuous contact episode. For $i \in \mathcal{S}_{\text{sep}}$, let c_i be the number of subsequent re-contact episodes, d_i their total duration, and T_i^{post} the post-separation horizon. We report

$$\text{ISR} = \frac{|\mathcal{S}_{\text{sep}}|}{|\mathcal{S}|}, \quad \text{RASR} = \frac{1}{|\mathcal{S}_{\text{sep}}|} \sum_i \mathbf{1}[c_i > 0], \quad \text{MREC} = \frac{1}{|\mathcal{S}_{\text{sep}}|} \sum_i c_i, \tag{32}$$

$$\text{MRD} = \frac{1}{|\mathcal{S}_{\text{sep}}|} \sum_i d_i, \qquad \text{MER} = \frac{1}{|\mathcal{S}_{\text{sep}}|} \sum_i \frac{d_i}{T_i^{\text{post}}}. \tag{33}$$

Thus RASR does not credit a rollout that never separated as "no re-contact," and MREC/MRD/MER retain separated rollouts with no re-contact as zeros. This conditioning is why the continuous diagnostics should not be interpreted as a simple $n/8$ re-contact count.

Contact conditioning. The audited Contact test bucket contains 209 scenes and 747 scene–time groups, but the physical post-impact comparison uses a frozen scene-disjoint lock of 8 anchors selected from its 747-target observation-legal pool. Each selected target is advanced under the exact nominal action a_0 until a simulator-observed contact state with at least 40 remaining evaluation steps is reached, and the resulting state fingerprint is then shared across methods before treatment begins. Consequently, the collision rate is 100%, and `post_contact_metric_eligible`=1 for every physical Contact row by construction. Overlap, penetration, and terminal-clearance quantities in Table 2 are therefore fully paired on these eight anchors; the continuous re-contact diagnostics in Table 3 follow the eligibility definitions above and are reported separately. The 100% contact conditioning should not be interpreted as a collision-avoidance score.

G Additional Qualitative Rollouts

All panels in this section use target-locked initial states and synchronized simulator times from the same frozen visualization pipeline as the quantitative evaluation. The five visualization candidates available for each stratum are arranged in a *post-hoc qualitative presentation order*: the rank is not an evaluation score and never enters the target lock, metric computation, or any table. We therefore retain only scenes that add distinct qualitative evidence rather than reproducing all fifteen candidates. Static keyframes are intended to expose geometric evolution; duration- and exposure-based claims remain those in Tables 1–3 and Table 6.

The supplementary archive `video.zip` contains the corresponding full-horizon renderings under `video/{safe,near,contact}/rank_XX/`. For every displayed rank, it includes an all-method synchronized video and an OC-RAP-versus-primary-comparator video. These videos use the same locked targets and rendering convention as the figures below and are supplied for qualitative inspection only; no numerical result is selected or computed from the videos.

G.1 Full Safe Comparison

The Safe stratum is designed to test non-interference, so its full external comparison is kept in the appendix to preserve main-text space while retaining the complete trade-off.

Table 6 isolates the intended Safe behavior. Both OC-RAP variants are the only rows with exact NUP $= 1$ and zero intervention. Diffusion Planner has the lowest collision (0.32%) and off-road (1.72%) rates, while PDM-Closed has the lowest isolated latency (163.24 ms). The evidence therefore supports nominal preservation, not metric-wise dominance.

G.2 Selected Qualitative Rollouts

Safe: non-interference. Figure 3 uses the highest-ranked Safe visualization candidate. Safe is intentionally the least dramatic regime: an effective recovery-aware planner should not introduce unnecessary corrective behavior when the nominal action remains well supported.

Table 6: Safe comparison on the common 174-scene observation-legal target lock. Lower collision, off-road, intervention, and latency are better; higher NUP is better. Latency is the mean steady-state deployed-planner time from isolated single-process/single-graphics-processing-unit (GPU) profiles.

Method	Coll.%↓	Off-road%↓	NUP↑	Interv.%↓	Latency (ms)↓
OC-RAP Balanced	0.84	1.93	**1.000**	**0.00**	234.95
OC-RAP Precision	1.02	2.01	**1.000**	**0.00**	227.84
GameFormer	1.15	2.30	0.919	44.66	208.76
PlanTF	2.87	2.30	0.939	34.58	191.82
PLUTO	1.08	1.85	0.606	98.66	193.92
PDM-Closed	1.21	2.83	0.804	31.31	**163.24**
PDM-Hybrid	1.17	2.87	0.812	31.31	163.79
IDM	2.30	4.02	0.879	92.04	168.46
Diffusion Planner	**0.32**	**1.72**	0.910	35.29	229.45

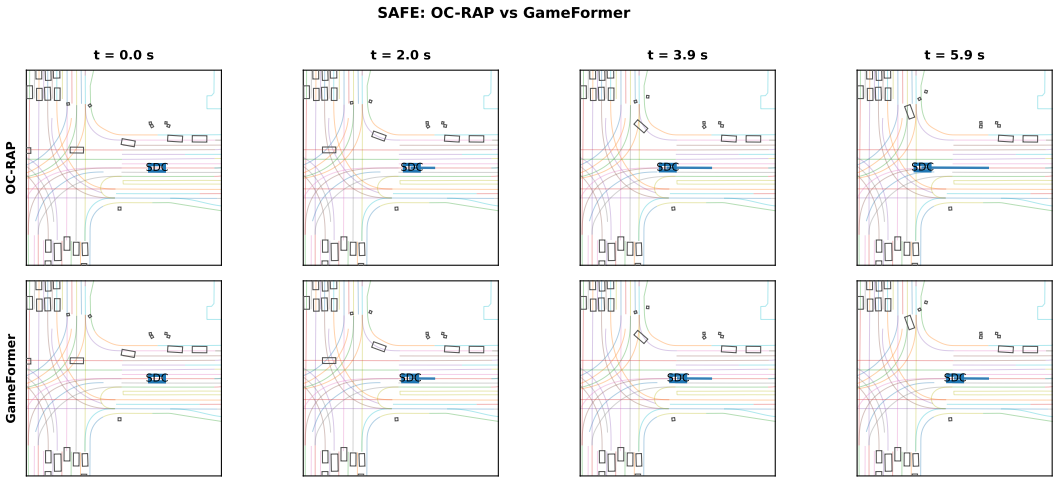


Figure 3: **Safe rank-01: OC-RAP vs. GameFormer.** Rows are methods and columns are synchronized keyframes. Both rollouts remain smooth and route-consistent over the displayed horizon; importantly, OC-RAP does not exhibit an unnecessary evasive correction merely because recovery machinery is available. This deliberately uneventful example is the qualitative counterpart of the Safe preservation result in Table 6: OC-RAP has zero intervention and NUP $= 1$ on the locked cohort. The single scene is illustrative and is not used to estimate those rates.

Near-Contact: lower-tail headroom. We use two complementary Near-Contact views. Figure 4 is a readable pairwise comparison against Plan-R1, whose non-collision clearance p05 (0.650 m) is the closest reported value to OC-RAP's 0.661 m. Figure 5 then exposes the same type of interaction across the complete audited comparator set.

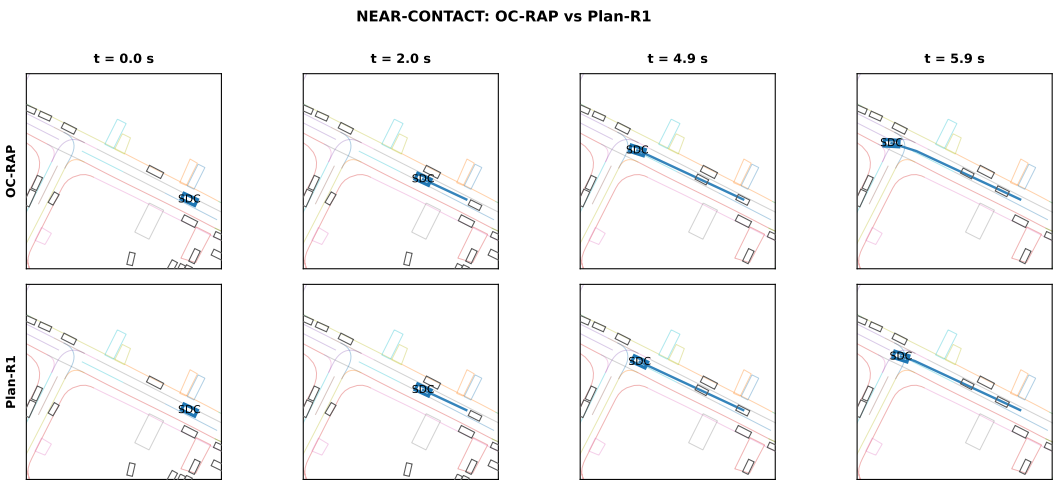


Figure 4: **Near-Contact rank-04: OC-RAP vs. Plan-R1.** This pair is included as a demanding geometric reference rather than a failure-only showcase: Plan-R1 is the closest listed comparator in non-collision clearance p05 (0.650 m versus 0.661 m for OC-RAP; Table 1). The synchronized frames keep both methods in the same scene and view, making the comparatively small differences in interaction geometry visible without changing the evaluation protocol. The lower-tail clearance and TTC conclusions remain full-cohort statements from Table 1.

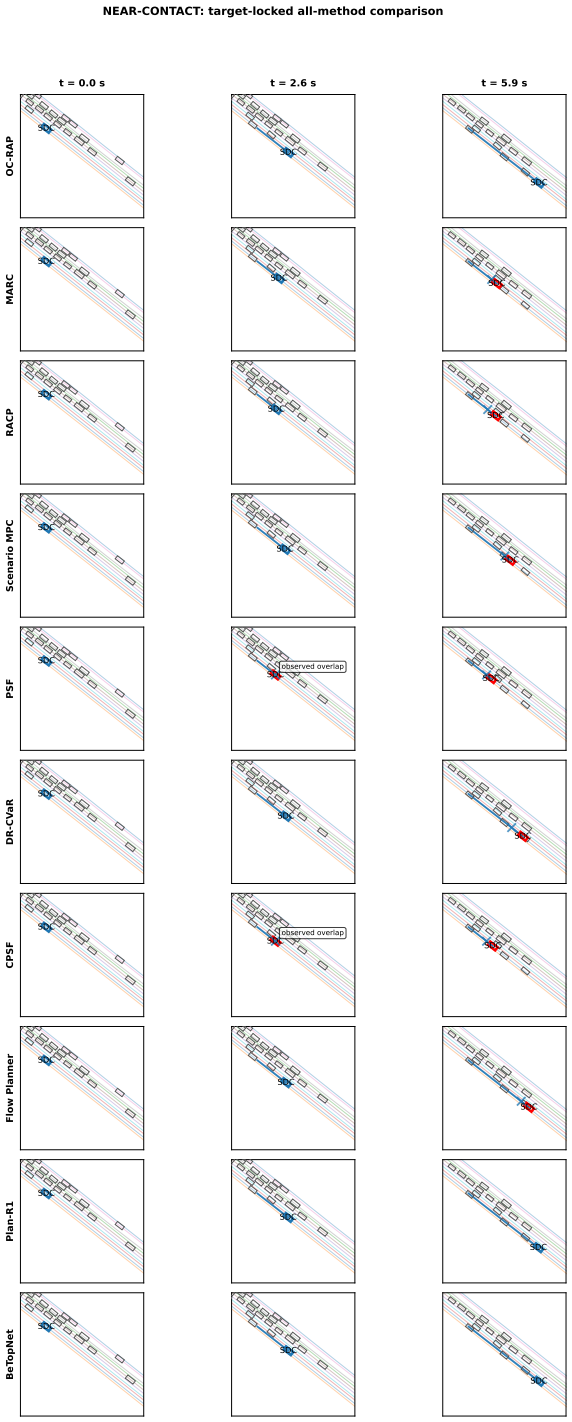


Figure 5: **Near-Contact rank-01 all-method comparison.** Rows contain OC-RAP and all nine Near-Contact comparators from Table 1; columns show synchronized start/critical/end keyframes. In this selected scene, the renderer explicitly marks a simulator-observed overlap for the Predictive Safety Filter and Conformal Predictive Safety Filter at the middle keyframe, while no such marker appears in the displayed OC-RAP frames. More broadly, methods that begin from the same target lock evolve to visibly different critical and terminal geometries. The panel is a qualitative stress-case cross-check only: collision, clearance, TTC, and exposure claims are computed over the full 250-scene cohort, not from this post-hoc visualization rank.

Contact: recovery debt repayment after a common observed-contact boundary. The Contact figures should be read differently from Near-Contact: every method begins from a shared simulator-observed contact state by protocol, so the question is how quickly and cleanly the planner separates afterward. We retain the rank-01 pair for a clear post-impact trajectory divergence, a rank-02 pair against the strongest external physical comparator in Table 2, and the rank-01 all-method panel for breadth.

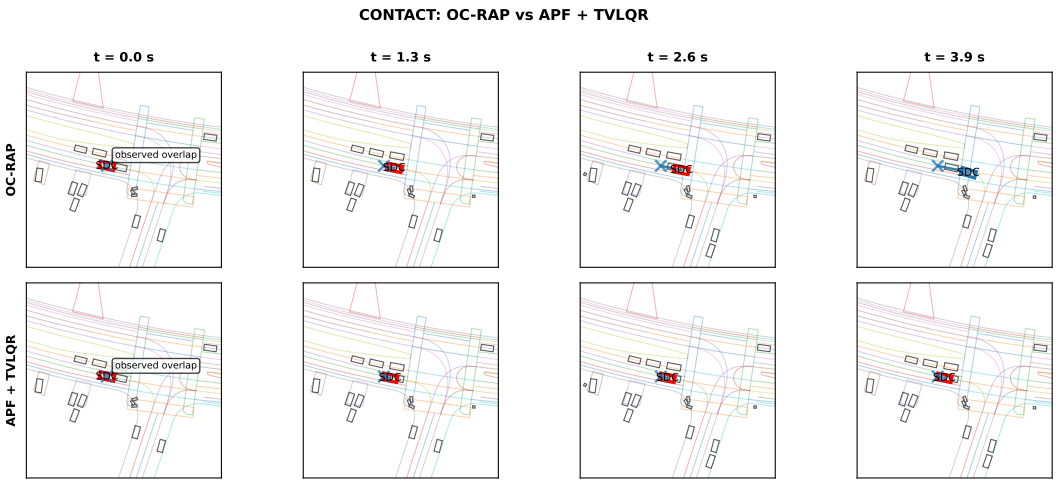


Figure 6: **Contact rank-01: OC-RAP vs. post-impact motion planning + TVLQR (labeled APF+TVLQR in the renderer, with APF denoting artificial potential field).** Both rows start from the same target-locked state with a simulator-observed overlap at $t = 0$, so this figure does not visualize collision avoidance. Instead, the synchronized post-impact frames show the recovery trajectories separating after the common boundary, providing a geometric view of the debt-repayment objective. Quantitative claims about overlap duration, penetration AUC, and terminal clearance use the fully paired eight-anchor results in Table 2; continuous re-contact claims use Table 3.

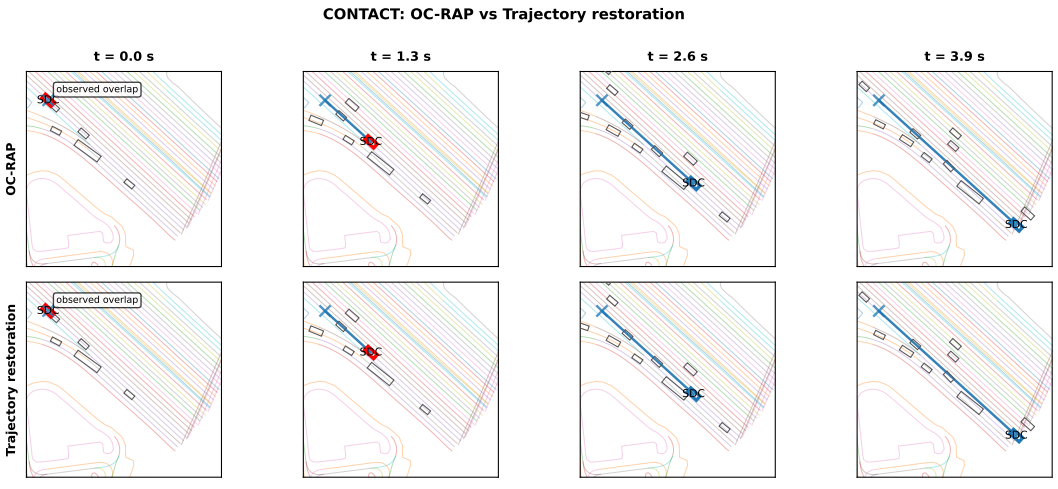


Figure 7: **Contact rank-02: OC-RAP vs. post-collision trajectory restoration.** Trajectory restoration is the strongest non-OC-RAP physical comparator in Table 2 for overlap duration, penetration AUC, and terminal clearance, making it a useful paired reference even though the route-level traces remain visually close. The shared overlap marker at $t = 0$ confirms identical contact conditioning; the static keyframes are not used to infer short-duration overlap or penetration differences, which are measured from the complete synchronized rollouts.

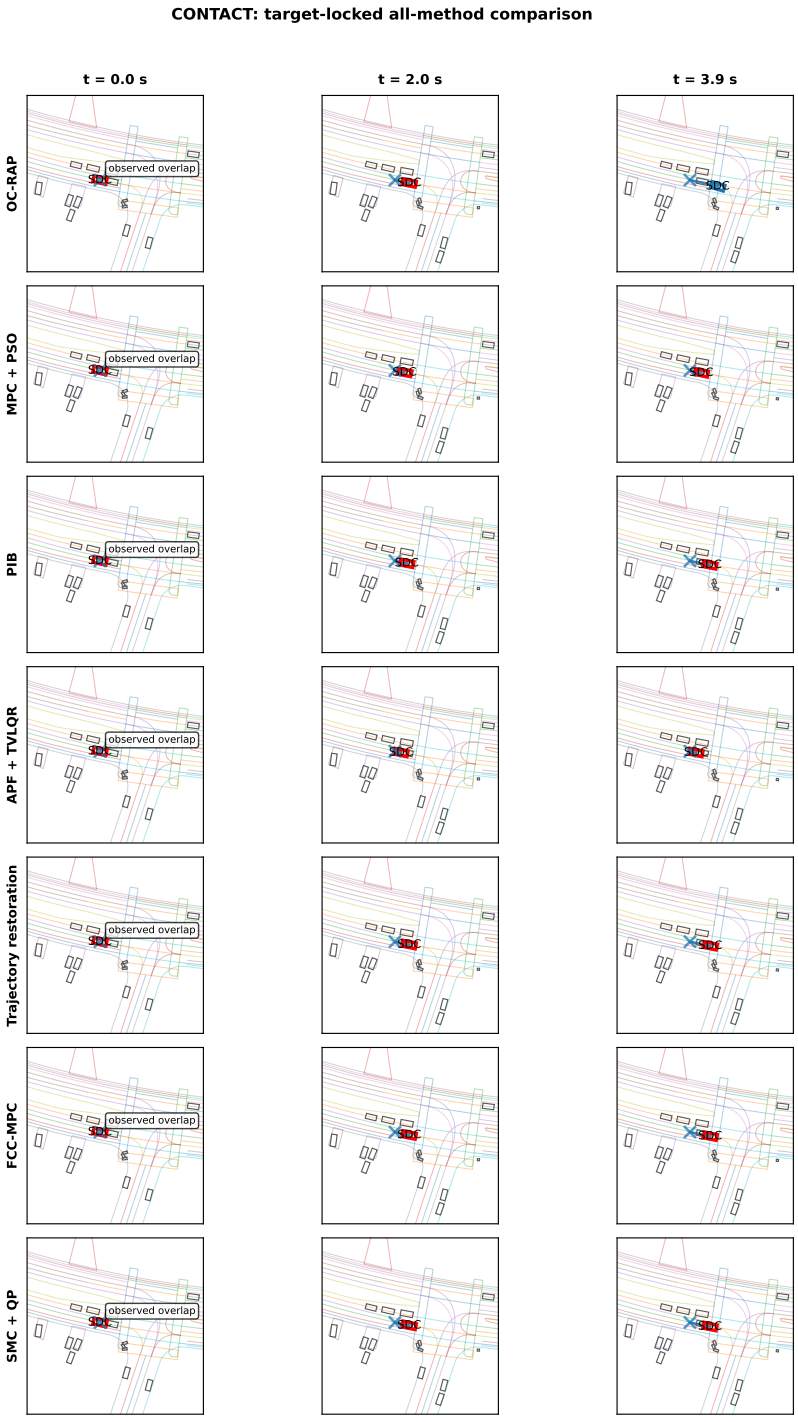


Figure 8: **Contact rank-01 all-method comparison.** OC-RAP and the six post-impact baselines from Table 2 are initialized from the same observed-contact anchor, with the common $t = 0$ overlap marker indicating this shared conditioning. The synchronized middle and terminal frames expose how recovery geometry diverges after that fixed boundary, giving a qualitative counterpart to overlap duration, penetration AUC, and terminal clearance; re-contact timing/exposure is quantified separately in Table 3. Because the Contact lock contains only eight fully paired anchors, this selected scene is presented as mechanism-level visualization rather than evidence about population collision frequency.